



## 저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

언론정보학석사학위논문

# 피드백과 목표가 크라우드소싱 결과물의 질에 미치는 영향

2015 년 8 월

서울대학교 대학원

언론정보학과

임 재 은

# 국 문 초 록

본 연구는 클라우드소싱 작업에서 커뮤니케이션을 통해 작업 결과물의 질을 높이는 방안을 탐구한다. 이전의 연구들은 보수의 양을 늘리거나 지급 방식을 달리하는 등 작업 결과물에만 초점을 맞추었다. 하지만 본 연구는 작업 자체보다는 작업자에 대해 관심을 갖고, 작업자와의 커뮤니케이션에 주목하였다. 선행연구를 통해 작업 관련 커뮤니케이션은 피드백과 목표로 나뉘며, 피드백은 긍정적, 부정적, 서술적 및 평가적 피드백으로 나뉘었다. 아울러 목표는 단계목표와 최종목표를 함께 주는 경우와 최종목표만 제시하는 경우, 표적형 목표와 성취형 목표 등으로 구분했다. 그리고 피드백과 목표의 효과에 대한 선행연구에 근거하여 가설을 설정하였다. 본 실험에 앞서 작업자와의 커뮤니케이션이 갖는 영향력을 알아보기 위해 예비연구를 실시했는데, 제3자가 결과물을 검토할 것이라는 가능성을 언급한 것 만으로도 작업자들은 작업에 더 많은 시간을 할애하였다. 그러나 작업의 질은 큰 차이가 없어 커뮤니케이션이 보다 구체적인 필요가 있었다. 이에 기반하여 본실험은 제시문의 길이를 줄이는 작업을 실험내용으로 하고, 4개의 피드백과 4개의 목표 등 8개의 보다 구체적인 커뮤니케이션을 작업자에게 제공하였다. 또한 보수의 양의 영향력과 커뮤니케이션의 영향력을 비교하기 위하여 두 가지 보수조건에서

실험을 진행했다. 이때 결과물의 질은 작업 결과물의 작업길이와 시간, 비용 등 3가지로 측정했다. 실험 결과 부정적 피드백을 제시했을 때 긍정적 피드백 조건과 같은 시간 및 비용을 들이고도 작업의 목적에 더 부합하는 결과물이 도출되어, 부정적 피드백이 결과물의 질적 향상에 효과적이었다. 또한 평가적 피드백이 서술적 피드백보다 작업길이 및 비용을 상대적으로 큰 폭으로 줄이고 있어, 평가적 피드백이 보다 효과적인 방식임을 알 수 있었다. 또 단계목표와 최종목표를 함께 제공하여 목표를 세부적으로 제시했을 때 결과물의 질은 최종목표만 제시되었을 때보다 향상되었다. 마지막으로 작업의 이유를 알려주는 성취형 목표가 체크리스트만 제공하는 표적형 목표보다 결과물의 질적 향상에 도움을 주었다. 위와 같은 연구결과에서 부정적 피드백, 평가적 피드백, 단계목표와 최종목표, 성취형 목표가 보다 효과적인 커뮤니케이션 방식임을 알 수 있었다. 그러나 보수의 양은 결과물의 질적 향상에 뚜렷한 영향력을 보이지 않고 커뮤니케이션 방식에 따른 차이가 더 뚜렷하여, 양질의 결과물을 도출하는데에는 커뮤니케이션이 더 효과적임을 알 수 있었다.

주요어: 클라우드소싱, 결과물의 질, 피드백, 목표, 아마존 메커니컬 터크

학 번: 2013-22844



# 목 차

1. 서론 .....	1
2. 선행연구 검토.....	6
2.1 기존 클라우드소싱 품질관리 연구 .....	6
2.2 작업흐름 디자인을 통한 클라우드소싱 품질관리 연구 .....	11
2.3 사회심리학의 반사실적 사고와 행동 연구 .....	14
2.4 피드백의 특성 및 효과 연구 .....	16
2.5 목표의 특성 및 효과 연구.....	21
3. 예비연구 .....	27
3.1 실험목적 .....	27
3.2 실험내용 .....	30
3.3 실험절차 .....	32
3.4 연구가설 .....	33
3.5 실험결과 .....	34
4. 연구문제 및 가설 .....	36
5. 연구방법 .....	42
5.1 참여자 모집 .....	42
5.2 실험내용 .....	43
5.3 실험절차 .....	45
5.4 실험처치 .....	50
5.5 측정 .....	53

6. 연구결과 .....	55
6.1 데이터 클리닝 .....	55
6.2 기술통계 .....	58
6.3 세부 연구결과 .....	61
6.3.1 커뮤니케이션과 보수의 양의 영향력 비교 .....	61
6.3.2 긍정적 피드백과 부정적 피드백의 효과 비교 .....	65
6.3.3 서술적 피드백과 평가적 피드백의 효과 비교 .....	73
6.3.4 단계목표+최종목표와 최종목표의 효과 비교 .....	83
6.3.5 표적형 목표와 성취형 목표의 효과 비교 .....	92
6.4 연구결과 요약 .....	101
7. 논의 및 시사점 .....	102
8. 결론 .....	110
9. 참 고 문 헌 .....	111
ABSTRACT .....	119

## 표 목차

표 1. 피드백의 유형 .....	19
표 2. 목표의 유형 .....	27
표 3. 예비연구 실험 조건 .....	28
표 4. 피드백 조건별 제공 메시지 .....	51
표 5. 목표 조건별 제공 메시지 .....	52
표 6. 키워드 리스트 .....	56
표 7. 조건별 참여자 수 .....	58
표 8. \$0.15 조건의 기술통계 .....	58
표 9. \$0.3 조건의 기술통계 .....	59
표 10. 통합 품질 기술통계 .....	63
표 11. 이원분산분석 결과표_ 커뮤니케이션 x 보수 .....	63
표 12. 회귀분석 결과 .....	64

## 그림 목차

그림 1. 어려운 작업 x 검토 조건의 작업 지시 화면.....	31
그림 2. 쉬운 작업 x 비검토 조건의 작업 지시 화면.....	31
그림 3. 분류작업 사진.....	32
그림 4. 예비연구_그룹별 작업시간 비교.....	35
그림 5. 예비연구_그룹별 작업 정확도 비교 .....	35
그림 6. 작업 가이드라인 .....	46
그림 7. 목표 조건에서의 사전 목표 화면.....	47
그림 8. 텍스트 단축 작업 화면 .....	47
그림 9. 커뮤니케이션 메시지 예시_표적형 목표 .....	48
그림 10. 긍정적 피드백과 부정적 피드백의 단계별 종속변인 .....	66
그림 11. 긍정적 피드백 vs. 부정적 피드백_종속변인 변화폭(\$0.15) .....	70
그림 12. 긍정적 피드백 vs. 부정적 피드백_종속변인 변화폭(\$0.3) .....	70
그림 13. 긍정적 피드백_보수조건별 종속변인 변화폭 .....	71
그림 14. 부정적 피드백_보수조건별 종속변인 변화폭 .....	71

그림 15. 서술적 피드백과 평가적 피드백의 단계별 종속변인 .....	74
그림 16. 서술적 피드백 vs. 평가적 피드백_종속변인 변화폭(\$0.15) .....	77
그림 17. 서술적 피드백 vs. 평가적 피드백_종속변인 변화폭(\$0.3) .....	77
그림 18. 서술적 피드백_ 보수조건별 종속변인 변화폭.....	80
그림 19. 평가적 피드백_ 보수조건별 종속변인 변화폭.....	81
그림 20. 단계목표+최종목표와 최종목표의 단계별 종속변인 .....	84
그림 21. 단계목표+최종목표 vs. 최종목표_종속변인 변화폭(\$0.15) .....	87
그림 22. 단계목표+최종목표 vs. 최종목표_종속변인 변화폭(\$0.3) .....	87
그림 23. 단계목표+최종목표_ 보수조건별 종속변인 변화폭 .....	89
그림 24. 최종목표_ 보수조건별 종속변인 변화폭 .....	90
그림 25. 표적형 목표와 성취형 목표의 단계별 종속변인 .....	93
그림 26. 표적형 목표 vs. 성취형 목표_ 종속변인 변화폭(\$0.15)	96
그림 27. 표적형 목표 vs. 성취형 목표_ 종속변인 변화폭(\$0.3) ..	97
그림 28. 표적형 목표_ 보수조건별 종속변인 변화폭 .....	98

그림 29. 성취형 목표_ 보수조건별 종속변인 변화폭 .....	99
-------------------------------------	----

# 1. 서론

클라우드소싱은 Howe(2006)가 아웃소싱에서 아이디어를 얻어 정의한 새로운 개념으로, ‘기존에 지정된 대리인 혹은 직원이 수행하던 일들을 누구인지 알 수 없는 다수의 사람들에게 자유로운 형태로 맡기는 행위’이다. 클라우드소싱으로 인해 사람들은 이전에 동원할 수 없던 많은 수의 사람들을 일시적으로 고용하고, 적은 비용으로도 빠른 시간 안에 작업을 완료할 수 있게 되었다. 수백 개의 이미지에 태그를 달아 구분하거나 동영상에 자막을 달기 위해 소리를 받아적는 등 소수의 사람들이 하기 어렵거나 힘든 작업들도 여러 명의 작업자들이 함께 일하여 빠른 시간 안에 완료할 수 있었다. 이처럼 클라우드소싱은 사람들이 원하는 많은 것들을 충족해주고 있으나, 결과의 질과 관련해서는 아직 부족한 점이 있다. ‘누구인지 알 수 없는 다수의 사람들에게 자유로운 형태로’ 일을 맡겼을 때 결과물의 질을 보장하기란 어렵기 때문이다. 따라서 결과물의 질을 높이기 위해 다양한 접근이 시도되고 있다. 가장 기본적으로 금전적 보상의 양과 제공방식에 따라 결과물의 질을 높이려는 시도가 있었고, 부적절한 작업 행위를 보이는 작업자들을 걸러 내거나, 투표를 통해 가장 좋은 결과물을 골라내기도 했다. 그러나 Mason 과 Suri(2012)의 연구에 따르면 보상의 양이 증가할 때 작업 결과물의 양은 증가하는 반면 질이

향상되지 못했다. Boner 등(2000)은 난이도를 고려하여 작업의 종류에 따라 인센티브를 지급하는 방식을 사용했하지만, 이 또한 일부 작업에만 효과를 나타냈다. 작업자들을 걸러내는 필터링 과정은 사후 작업이므로 시스템 전반에 걸쳐 지속적으로 질적 향상을 추구하기는 어려웠다. 투표 방식도 유사한 이유로 데이터 수집 과정에서 품질을 관리하기 어렵고, 작업 성격에 따라 투표 방식이 적합하지 않은 경우도 있어 한계를 보였다.

위와 같이 기존 방식은 클라우드소싱에서 작업의 질적 향상에 제한된 영향력만을 가지고 있다. 이러한 한계는 작업 자체에 집중하면서 작업자에게 동기를 제공하는 방식 혹은 이와 관련된 요소들에 대해 간과했기 때문에 발생했다. 우선 기존 방식들은 작업자의 경제적 동기 외의 동기나 작업자의 속성을 고려하지는 못했다. 물론 금전적 동기는 클라우드소싱 작업자에게 중요하다. Ross 등(2010)의 연구에 따르면 메커니컬 터크의 작업자 중 상당수가 인도인이며, 이들은 클라우드소싱을 2차적 혹은 1차적 수입원으로 생각하고 있어 금전적 보상의 영향력을 무시할 수는 없다. 그러나 동시에 금전적 동기만으로 작업자를 이해하는 것은 지나치게 제한된 시각이다. 사람은 다양한 동기를 가지고 일하기 때문이다. 설사 경제적 동기가 참여로 이끄는 동기 중 1순위라고 하더라도, 클라우드소싱에서 지급하는 보수의 수준은 일반적 보수보다 낮고, 클라우드소싱 사이트 내에서 관습적으로 일정



수준에 맞춰져 있다.<sup>1</sup> 따라서 보수의 양을 계속 늘려 어느 수준을 넘어서게 되면 실제 클라우드소싱 환경에 적용하기는 어려울 수 있다. 하지만 경제적 동기에만 집중하는 것이 아니라 이 동기를 다른 방식으로 활용할 수 있다면, 더욱 좋은 결과를 낼 수 있을 것이다. 경제적 동기가 큰 작업자라면 자신의 작업 결과가 의뢰자의 맘에 들지 않아 보수를 지급받지 못하는 상황에 대한 불안감이 클 것이다. 이 때 최종 제출 전 자신의 답안에 대해 누군가 검토해주고 보완해야 할 점 등을 알려줄 수 있다면 작업자의 불안은 해소될 수 있고, 해당 작업자의 경제적 동기를 인위적으로 높이지 않고도 충분히 좋은 결과물을 낼 수 있을 것이다.

또한 작업 자체에만 주목하여 사후 처리 위주로 결과물의 질을 높이려한 시도도 작업자의 동기를 충분히 고려하지 못했다. 주로 사용된 필터링 방식은 결과가 모두 나온 후에 부적절한 작업자를 제외하는 방식으로, 작업 전 혹은 작업 중에 동기를 자극할 수 없기 때문에 품질 향상에 한계를 가지고 있었다. 이러한 점을 보완하기 위해 Dow 등(2011)은 양치기가 양을 몰고 가는 것처럼 한 명의 관리자가 다수의 작업자들을 담당하는 시스템(Shepherding system)을 제안하였다. 그러나 1명 혹은 소수의 관리자가 결과물을

---

<sup>1</sup> 가장 낮은 보수는 \$0.01부터 시작하며, 보통 \$0.03~\$0.1 정도로 지급된다.

읽고 검토하여 피드백을 제공하는 이 시스템은 소수의 관리자에게 지나친 부담을 줄 수 있고, 클라우드소싱보다는 전통적인 작업 방식에 가까워 클라우드소싱 시스템을 개선하는 방안으로서는 한계가 있다.

하지만 작업 결과물에 대해 작업자와 직접 커뮤니케이션 할 수 있다면, 필터링과 양치기(Shepherding) 방식의 한계를 극복할 수 있을 것이다. 결과물을 일정한 기준에 따라 자동으로 분석하고 분석 결과를 작업자에게 알려준다면, 이전에는 필터링 대상이 되었을 부적절한 결과물도 적절한 결과물로 바뀔 수 있다. 또 결과물의 전반적인 품질 향상도 기대할 수 있는데, 필터링 대상이 아니었던 결과물에 대해서도 작업자가 다시 한 번 검토할 기회가 있어 품질이 높아질 수 있기 때문이다. 아울러 이처럼 자동화된 커뮤니케이션은 사후 처리 방식의 한계로 지적되었던 관리자의 부담이나 클라우드소싱 환경과의 부적합성 등을 개선할 수 있다.

이처럼 결과물의 질을 향상시키는 일과 관련된 기존 연구들은 한계를 보였으나, 그러한 한계들은 작업자와의 커뮤니케이션을 통해 개선될 가능성이 있다. 따라서 본 연구는 기존 방식의 한계를 극복하고, 작업자의 동기를 보호하고 자극하는 방안으로 작업자와의 커뮤니케이션에 주목했다. 개별 작업자 모두를 대상으로 특정한 동기가 높아지도록 일괄적으로 자극을 가하는 것이 기존 방식이었다면, 여기서는 작업자가 어떤 동기를 가지고

있는 상관없이 이들과 소통함으로써 각자의 동기를 통해 작업 결과물의 질을 높이하고자 한다.

클라우드소싱은 일을 맡기는 사람과 하는 사람이 서로를 알지 못하며, 서로간 커뮤니케이션이 극히 제한되어 있다. 그래서 작업 설명을 잘 써놓더라도 작업자 입장에서는 자신의 결과물이 설명과 맞는지, 의뢰한 측의 의도는 정확히 무엇인지 등 궁금한 점이 발생할 수 있다. 물론 의뢰자의 이메일을 통해 커뮤니케이션이 가능하기는 하지만, 이 또한 충분치 못하다. 적은 보수를 받고 단순작업을 하는 작업자에게 의뢰자와 이메일로 질의응답을 거친 다음 다시 작업을 하는 만큼의 열의와 성의를 기대하기에는 무리가 있기 때문이다. 이처럼 클라우드소싱에서는 작업과 관련된 커뮤니케이션의 간극이 발생하며 그 간극이 작업 결과물의 질에 악영향을 줄 수 있으나, 작업자 혹은 의뢰자 개인의 노력으로 이를 개선하는 것은 어렵다.

본 연구는 이와 같은 클라우드소싱 내 작업자와 의뢰자 간 커뮤니케이션의 간극을 해결하기 위해 현재 상황이 어떤지 알려주는 방식과 앞으로 무엇을 해야 하는지 제시해주는 방식 등 2가지 커뮤니케이션에 주목했다. 피드백을 제공하는 경우는 전자에 해당하며, 기술적이고 설명적이어서 자신의 결과물에서 보완하거나 개선해야 할 점을 스스로 판단할 수 있다. 또 자신의 결과물에 대한 피드백이므로 작업자의 참여도를 높일 수도 있다. 그러나

구체적인 선택지는 제공되지 않아 작업자의 특성이나 작업의 성격에 따라 효과는 달라질 수 있다. 반면 앞으로 무엇을 해야 하는지를 알려주기 위해 목표를 제공할 경우, 작업자는 구체적인 가이드라인을 가지게 되므로 결과물의 질을 높일 수 있으나 모든 작업자에게 동일하게 제공되기 때문에 작업자가 피드백만큼의 관심을 갖지 않을 수도 있다.

이처럼 개별 작업자의 동기를 유지 또는 강화하여 결과적으로 클라우드소싱 결과의 질을 높이는 방식으로 피드백과 목표는 각각 장점과 단점을 가지고 있다. 본 연구는 피드백과 목표를 클라우드소싱 작업자에게 동기를 부여하기 위한 커뮤니케이션 방식으로 보고, 이들이 결과물의 질적 향상에 미치는 영향력을 탐구하였다. 이를 위해 피드백과 목표에 대한 선행연구를 검토하여 각 방식의 세부 유형 및 그 효과를 파악하고, 클라우드소싱 환경에 적합한 방식을 선별하여 연구를 진행했다.

## 2. 선행연구 검토

### 2.1 기존 클라우드소싱 품질관리 연구

그동안 클라우드소싱 결과의 질을 유지, 관리 및 향상시키기 위하여 시도된 방식은 금전적 보상, 필터링, 주식 등 3가지로 정리할 수 있다. 먼저 금전적 보상은 클라우드소싱에서 가장

보편적으로 이용되는 방식이므로 많이 연구되었는데, 특히 아마존사의 메커니컬 터크(Amazon's Mechanical Turk)가 그 대상이 되는 경우가 많았다. 메커니컬 터크는 데이터를 얻고 싶은 의뢰자(requester)와 일을 하고 보수를 얻으려는 작업자를 연결하는 온라인 마켓이다. 의뢰자는 자신이 원하는 대로 HIT(Human Intelligence Task)라고 하는 작업 한 단위를 디자인한 후 사이트에 등록한다. 그리고 작업자는 의뢰자들의 작업들이 올라온 게시판에서 작업에 대한 설명을 읽거나 보수의 양을 검색하면서 작업을 고른다. 작업자가 스스로 선택한 작업에 참여한 다음 결과물을 제출하면 작업자는 의뢰자로부터 약속한 금액을 받는다. 메커니컬 터크의 이러한 흐름을 활용하여 연구자들은 금전적 보상의 양과 방법(scheme)에 따라 결과의 질을 높이하고자 시도하였다.

그러나 금전적 보상의 양은 높은 질을 보장해주지 못하고 있었다. Mason 과 Watts(2009)의 연구에서 같은 일에 대해 보상을 차등적으로 주었을 때 유의미한 품질의 차이는 보이지 않았다. 이는 Mason 과 Suri(2010)의 연구에서도 유사하게 나타났는데, 보상의 양이 증가할 때 사람들이 완료한 작업의 수, 즉 결과물의 양은 증가하였으나 질이 향상되지는 않았다. 또한 보상의 방식에 대한 연구도 진행되어 Boner 등(2000)은 인센티브 지급의 효과를 높이기 위해서는 작업의 난이도를 고려하여 작업의 종류에 따라

주어야 한다고 주장했다. Kim 등(2012)은 보상의 방식을 단위비율(piece-rate), 할당(quota), 처벌(punishment) 3가지로 나누어 분석했다. 단위비율은 작업자가 완성한 작업 수에 따라서 보수를 지급하는 방식으로 많이 작업한 만큼 많이 받을 수 있어 동기를 자극하는데 효과적이나, 한계효용이 감소하는 것에 취약하다. 할당은 일정한 목표를 설정해주는 것으로, 작업자에게 할당된 모든 작업을 완료해야만 보수를 지급한다. Boner 등(2000)은 이러한 방식이 복잡하고 어려운 작업에서 가장 효과적이라고 했으며, Mason 과 Watts(2009)도 단어퍼즐 맞추기 작업에서 할당 방식의 효과를 검증했다. 그러나 이러한 할당 방식은 크라우드소싱에서는 낯선 방식이어서, 크라우드소싱 전체에 적용하기에는 무리가 따를 수 있다. 마지막으로 처벌 방식은 보수를 지급하지 않거나 원래 보수에서 일정 비율을 제한 나머지를 지급하는 방식이다. Shaw 등(2011)은 작업자에게 작업 결과가 나머지 작업자들의 결과와 일치하지 않을 경우 원래 금액에서 10%를 제한 보수를 제공하겠다는 처벌 규정을 설정한 실험이 품질 향상에 가장 효과적이었다고 설명했다. 그러나 나머지 작업자들의 결과가 항상 옳다고 볼 수는 없으며, 단순 작업 외에 아이디어를 수집하는 등 다양성이 요구되는 작업에는 적용할 수 없어 여전히 한계를 가지고 있다.

금전적 보상과 함께 ‘불량한’ 작업자를 골라내는 필터링 방식도 품질관리를 위해 자주 사용되는 방식이다. 대부분 자동화된 방식으로 크라우드소싱 환경을 유지하고 관리하는 것을 목적으로 하고 있었다. Whitehill 등(2009)은 부적절한 작업자들을 스팸머(spammer)라고 규정하며, 스팸머를 찾아내어 사전에 차단할 수 있는 방식을 연구하였다. Kim 등(2012)은 단순히 보수만을 목적으로 품질이 낮은 결과를 양산하는 작업자들을 랜덤 클릭어(random clicker)라고 지칭하고 이들을 필터링하기 위한 방식을 고민하였다. 이들은 피어슨의 카이제곱 검정을 사용하여 작업 결과가 무작위적인 데이터인지 아닌지 가려냈고, 랜덤 클릭어가 배제된 상태인 필터링 된 결과는 실험실 결과와 유사한 것으로 나타났다. 필터링은 부정적인 결과를 차단할 수 있고, 자동화가 용이하다는 장점을 가지고 있다. 그러나 크라우드소싱 환경에 확대 적용하는데는 다소 무리가 있다. Kochhar 와 Mazzocchi(2010)는 필터링 방식을 사용할 때 걸려져야 할 부적합한 사례는 정의되어 있으나, 적합한 결과물에 대해서는 명확하게 정의되어 있지 않다는 점을 지적했다. 작업자들이 만들어낸 결과물 중 부적합한 사례를 제거한다면 처음 상태에 비해 품질이 향상될 수 있으나, 그 이상으로 품질을 향상하기 위해서는 적합한 사례에 대한 정의도 필요하기 때문이다.

마지막으로 주석(Annotations)은 자동화 방식으로 판명하기 어려운 데이터에 대해 여러 명의 작업자들이 검토하는 방식이다. 비전문가라도 여러 명이 검토할 경우에는 그 결과가 신뢰할 만하다는 사실은 Snow 등(2008)의 연구에서 입증되었다. 이들은 5개의 자연어 작업을 제시하고 비전문가와 전문가의 주석 결과 간의 질적 차이를 알아보았는데, 적은 수의 비전문가라도 전문가만큼의 결과를 낼 수 있다고 주장하였다. 이러한 결과를 바탕으로 인간의 언어적 능력을 활용한 다양한 연구가 진행되었다. 트위터를 크롤링한 결과 만들어진 방대한 데이터에서 사람들의 이름만을 추출하는 일은 시간과 비용이 많이 걸리지만, Finin 등(2010)은 클라우드소싱을 활용하여 이름들을 효과적으로 추출했다. 이처럼 데이터를 찾아내는 작업 외에도 자연어 처리 과정에서 주석 방식은 결과물의 질을 높이기 위해 활용되었다. Akkaya 등(2010)은 클라우드소싱을 통해서 단어 의미의 주관적 사용을 판단하는 실험을 하였다. 같은 단어라도 사용하는 맥락에 따라 의미가 달라지거나 애매해질 수 있는데, 이 경우 자동화로 판별해내는데 한계가 있어 클라우드소싱의 많은 작업자들에게 투표를 하게 했다. 그 결과 신뢰할 만한 데이터가 추출되어 주석 방식의 효과가 검증되었다. 이처럼 주석은 자동화가 어려워 사람의 지능이 요구되는 작업에 적용할 수 있고, 클라우드소싱의 장점을



잘 활용할 수 있다. 다만 명확한 정답이 있어 다수의 의견일수록 정답에 가까워지는 성격이 아닌 작업일 경우에는 적용하기 어렵다.

위와 같이 금전적 보상, 필터링, 주석 등 기존 연구에서 주로 연구된 3가지 방법은 클라우드소싱 내 결과물의 질을 향상하는데 기여하였으나 여전히 한계를 가지고 있었다. 금전적 보상은 한계 효용이 줄어들면서 어느 수준 이상의 동기 부여가 어려웠고, 필터링의 경우 필터링 대상이 아닌 작업 결과물에 대한 정의가 부정확하다는 점이 한계로 지적되었다. 또한 주석의 경우 다수의 답이 정답일 확률이 높은 작업에만 적용할 수 있었다. 그리고 3가지 방법 모두 다양한 작업에 적용하기 어렵고, 그 효과가 단기적이라는 공통적인 한계가 있었다. 그러므로 다양한 작업에 적용할 수 있도록 보편적이면서도 장기적인 효과를 기대할 수 있는 방식이 필요하다는 점을 도출할 수 있었다. 최근 이러한 한계를 보완하기 위한 새로운 방안으로 작업흐름 디자인이 주목받고 있다.

## 2.2 작업흐름 디자인을 통한 클라우드소싱 품질관리 연구

과거에 클라우드소싱은 영수증 사진을 보고 표에 그대로 베끼는 등 단순한 작업을 위해 활용되는 경우가 많았다. 하지만 점차 문서 작업이나 교정 등 복잡한 작업까지 클라우드소싱을 통해

이루어지고 있고, 학계에서도 이처럼 복잡한 작업을 클라우드소싱에서 진행할 때 어떻게 작업의 질을 높일 수 있는지 고민하고 있다. 그 중 하나로 등장한 것이 작업흐름 디자인으로, 복잡한 작업을 여러 단계로 나누어 진행하는 방식을 일컫는다. 관련하여 Bernstein 등(2010)은 ‘Find-Fix-Verify’라는 3가지 패턴을 디자인해 자연어 처리방식에 적용했다. 이 패턴은 품질 향상이 필요한 부분들을 찾고(Find), 이를 교정하거나 요약하는 등의 작업(Fix)을 거친 결과들을 모은 후, 마지막으로 이 중 부적합한 결과를 걸러내는 과정(Verify)을 거친다. 이러한 작업흐름 디자인 방식은 복잡한 작업을 그대로 클라우드소싱에 의뢰할 경우 원하는 결과를 얻기 어렵기 때문에 등장하였다. 게으른 작업자들은 가장 쉽고 간단한 작업만 하고 보수를 받고자 할 것이고, 지나치게 의욕적인 작업자들은 과도하게 수정하여 역시 부적절한 결과를 낼 것이기 때문이다.

Kittur(2011)도 유사한 관점에서 클라우드소싱 환경에서 자연어 처리를 위한 프레임 워크를 디자인하였다. 그는 어렵고 복잡한 과제를 내고, 참여(Participation), 지도(Map), 단축(Reduce) 등 3가지 단계를 가진 크라우드포지(CrowdForge) 프레임이 얼마나 결과물을 질적으로 향상시킬 수 있는지 알아보고자 했다. 크라우드포지를 통해 만들어진 결과물을 전문기자, 전문 학술지 저자, 소셜 컴퓨팅을 연구하는 대학원생 등에게 보여주고 평가했을

때 그 결과는 상당한 수준 이상이였다. Bernstein 과 그의 동료들, 그리고 Kittur 의 연구를 볼 때 복잡한 작업도 여러 단계로 나누어 진행될 수 있도록 작업 환경을 디자인한다면 충분히 결과의 질을 향상할 수 있을 것으로 보인다.

그러나 아직까지 해결되지 못한 점이 있는데, 작업흐름 디자인 방식에서도 여전히 작업자들이 작업에 관한 정보를 얻고 문제점을 해결하는 과정에서 필요한 커뮤니케이션이 부재하다는 점이다. 보수나 사후 필터링 등 작업 외의 요소에 집중한 기존 방식과 달리 작업흐름 디자인은 제공 방식에 집중했지만, 여전히 작업자와의 충분한 의사소통이 어렵다. 특히 텍스트 다듬기처럼 복잡한 작업의 경우 고려해야 할 점이 많은데, 이에 대한 커뮤니케이션이 부족하기 때문에 작업자들의 능력에 기대게 된다. 의욕이 높고 능력이 좋은 작업자만을 고용할 수 있다면 좋은 결과를 낼 수 있겠지만, 그렇지 않을 경우에 결과물의 질을 보장하기는 어렵다. 따라서 앞선 연구들에 더해 작업흐름 디자인을 보완하여 더욱 장기적이고 효과적으로, 그리고 작업자들의 동기와 능력을 이끌어낼 수 있는 방안이 필요하다.

크라우드소싱은 ‘다수의 지혜(Wisdom of the crowd)’라는 가치를 내걸고 시작되었으나, 지금까지는 ‘다수의 시간’ 혹은 ‘다수의 작업’이었을 뿐 지혜를 활용하기는 어려웠다. 기본 가치를 실현하기 위해서는 다수의 지혜를 활용할 수 있도록 자신의 작업이

어떻게 진행되고 있으며 더 좋은 결과를 만들기 위해서 어떤 점을 더 활용해야 하는지 등을 알려주는 환경이 필요하다. 즉, 작업 관련 커뮤니케이션을 활성화해야 한다. 작업 관련 커뮤니케이션은 단순히 작업자에게 어떤 정보를 제공하는 것 뿐만 아니라, 작업자의 행동에 영향을 줄 수 있기 때문이다. 본 연구는 그 이유를 사회심리학의 반사실적 사고 연구에서 찾았다.

## 2.3 사회심리학의 반사실적 사고와 행동 연구

반사실적 사고(counterfactual thinking)는 과거에 일어난 일이나 상태에 대한 대안의 정신적 표상(mental representation)으로, 과거에 일어난 일에 대해 ‘만약에 ~했다면’ 등의 방식으로 생각하는 것이다(Smallman 과 Roese, 2009). 이와 같은 사고는 행동교정 기능을 가지고 있어 행동 변화 및 성과 개선을 이끌어낸다(Roese, 1994). Epstude 와 Roese(2008)의 연구에 따르면 목표와 연계된 행동에서 반사실적 사고는 목표와 관련된 행동을 알려줌으로써 목표를 달성하는데 도움을 줄 수 있다. 이러한 영향은 내용 중립적 반사실적 사고와 내용 관련 반사실적 사고 등 2가지로 설명될 수 있다. 먼저 내용 중립적(content-neutral) 반사실적 사고는 맥락에 관계없이 대안을 고려하는 방식으로, 실패에 대한 반성 및 과거에 대한 시뮬레이션 등을 포함한다. 따라서 내용 중립적 방식은 반사실적

사고방식(counterfactual mindset)을 형성하여 행위자에게 특정 행동뿐만 아니라 관련된 다른 행위에까지 영향을 미치는 동기를 부여할 수 있다. 반면 내용 관련(content-specific) 반사실적 사고는 특정한 행동에 대한 의도를 형성하여 관련 행동에 영향을 주는데 집중한다. Bargh 등(2001)에 따르면 내용 관련 반사실적 사고는 부정적 상황에서 행위의도를 형성하고 행동으로 이어지기 때문에 전반적인 동기부여를 가져오는 내용 중립적 방식과는 다르다.

Smallman 과 Roesse(2009)는 이러한 반사실적 사고가 행동에 미치는 영향을 3가지 실험을 통해 탐구하였다. 먼저 이들은 반사실적 사고를 할 때 행동 판단에 걸리는 시간이 짧아진 결과를 토대로 반사실적 사고가 행동 관련 판단을 증진시킨다고 밝혔다. 두 번째 실험에서는 행위 의도가 내용 중립적 방식의 반사실적 사고가 동기를 부여한 결과인지 혹은 내용 관련 반사실적 사고의 교정적 과정 결과인지 알아보고자 했다. 그 결과 제시된 문제와 연관된 정보를 제공받은 사람들이 그렇지 않은 사람들보다 빠르게 판단하여, 내용 관련 반사실적 사고가 행동 의도에 영향을 주는 것으로 나타났다. 세 번째 실험에서는 반사실적 사고가 행위 의도와 무관한 판단보다는 행위 의도에 더 큰 영향을 주는 것으로 나타났다.

위의 3가지 실험을 통해 반사실적 사고는 행위자가 내용과 관련된 정보를 이용하여 특정 행위에 대한 의도를 형성하고

판단하게 한다는 점이 밝혀졌다. 즉 과거 행동에 대해 정보를 제공하고 미래의 행위의도를 형성하고자 한다면, 특정 행위와 연계된 자세한 정보를 제공하여 반사실적 사고방식의 교정적 과정을 거치도록 해야 보다 효과적이라는 점을 알 수 있다. 따라서 작업자들에게 과거의 작업결과를 보여주고 ‘열심히 일하자’라는 메시지를 제공하는 방식으로 동기를 부여하는 데는 한계가 있다. 부정적인 작업 결과를 긍정적으로 바꾸기 위해서는 해당 행동에 집중하여 행동과 관련한 자세한 정보를 제공하되, 반사실적 사고를 유도하여 행위 의도에 직접적으로 영향을 주는 것이 더 효과적일 것이다. 정보를 제공하는 방식은 개별 작업자의 결과물에 집중하는 방식인 피드백이거나, 작업 과정에서 달성해야 할 구체적인 목표를 주는 방식일 수도 있다. 이러한 이유로 다음 단락부터는 피드백과 목표와 관련된 연구를 살펴보고, 클라우드소싱에 어떻게 적용할 수 있는지 검토하고자 한다.

## 2.4 피드백의 특성 및 효과 연구

피드백과 관련된 연구는 피드백의 특성을 구분하고, 그에 따른 효과를 검증하는 방식으로 이루어졌다. 이차연과 박영석(2006)은 Jones 와 Gerard(1967)의 논의에 근거하여 업무수행 피드백 특성을 구분하였다. 사회심리학자인 Jones 와 Gerard 는 개인은 자신의 요구를 충족하기 위해 혹은 환경에 대한 정보를 얻고 이해하기

위해 피드백에 의존한다고 주장했다. 이차연과 박영석은 이들의 연구를 인용하여 일 처리 과정에 초점을 두고 개인의 행동에 대해 구체적으로 서술하는 정보 피드백(information feedback)과 결과에 초점을 두고 평가하는 내용을 담은 보상 피드백(reward feedback)으로 업무수행 피드백을 나누었다. 이러한 구분은 Harackiewicz 등(1987)의 연구와 연결될 수 있는데, 이들은 경쟁을 강조하거나 결과를 평가하는 피드백과 과정에 대한 정보를 서술하는 피드백을 비교했다. 평가에 대한 압박감이 동기 유발에 부정적인 영향을 줄 수 있기 때문에, 이들은 정보 위주의 피드백에 초점을 맞출 때 작업자의 동기를 유지하고 강화할 수 있다고 설명했다. 두 연구를 종합해보면 피드백의 특성에 따라 정보 중심적 피드백과 보상 중심적 피드백, 평가적인 피드백과 서술적인 피드백이라는 유형을 도출할 수 있다.

준거기준에 따라서 피드백을 분류하는 연구도 있었다. 권은주와 김성일(2003)의 연구는 수행의 질이나 과제에 대한 흥미를 높이기 위해서는 타인과의 비교보다는 학습자 본인 혹은 과제 자체에 초점을 두는 피드백을 제공해야 한다고 주장했다. Gallien 과 Oomen-Early(2008)의 연구에서는 개인적 피드백(personalized feedback)과 종합적 피드백(collective feedback)을 구분하여, 각 유형 간 학생들의 만족도와 학업 성취도, 그리고 그들이 인지하는 교수자와의 연결성을 비교하였다. 개인적 피드백 조건에서는

학생들이 작성한 과제 각각에 대해 개별적인 피드백을 제공하였다. 집단형 피드백 조건에서도 각 과제에 피드백을 제공하였으나, 그 내용이 학생 개인에 대한 피드백이 아니라 학급 전체에 대한 평가인 점이 달랐다. 이러한 연구들을 종합해볼 때 준거기준에 따라서 피드백은 상대적 피드백과 절대적 피드백으로 구분된다.

피드백을 제공하는 시기로 피드백의 특성을 구분한 Kulik 와 Kulik(1988)의 연구는 따르면 결과가 나온 즉시 제공하는 즉각적 피드백(immediate feedback)과 모든 결과가 나온 후 최종적으로 제공되는 지연된 피드백(delayed feedback)으로 피드백을 구분했다. 연구 결과 단순반복 작업의 경우 지연된 피드백이 효과적이나 복잡하고 고차원적 사고를 요하는 경우 즉각적인 피드백이 더 효과적이었다. 아울러 한주희와 고수일(2005)은 피드백의 전달 측면에 주목했다. 이들은 집단 작업환경에서 피드백의 효과를 살펴보았는데, 피드백이 전달하는 측면에 따라 긍정적 피드백과 부정적 피드백으로 구분하였다. 연구 결과 긍정적 피드백 뿐만 아니라 부정적인 피드백을 받은 경우에도 팀 몰입과 팀 혁신노력이 증가한 반면, 피드백이 주어지지 않은 경우에는 팀 몰입이 유의미하게 감소하였다. 따라서 피드백이 팀 단위 작업활동에 유의미한 영향력을 가지고 있으며, 이는 긍정적 피드백 뿐만 아니라 부정적 피드백에서도 나타난다는 사실을 알 수 있다. 위와



같이 선행연구들을 통해 다양한 피드백의 특성 및 효과를 파악할 수 있었는데, 이를 정리한 결과는 아래 표와 같다.

표 1. 피드백의 유형

구분	유형	
목적	정보피드백	작업 과정 및 행동 관련 정보를 얻어 환경을 이해함
	보상피드백	결과에 초점을 두고 평가하여 보상을 얻음
기술대상	평가적 피드백	경쟁을 강조하고 결과에 대한 평가를 전달함
	서술적 피드백	작업 과정에 대한 정보를 전달함
준거기준	상대적 피드백	사회적 준거에 따라 평가함
	절대적 피드백	절대적, 개인적 준거에 따라 평가함
제공시기	즉각적 피드백	개별 결과가 나온 즉시 피드백을 제공함
	지연된 피드백	모든 결과가 나온 다음 피드백을 제공함
전달측면	긍정적 피드백	결과물의 긍정적인 측면을 집중적으로 전달함
	부정적 피드백	결과물의 부정적인 측면을 집중적으로 전달함

이와 같이 피드백의 특성 및 그 효과에 대해 살펴본 결과 피드백은 그 특성에 따라 총 5가지 유형으로 분류되었다. 각 피드백 모두 그 효과가 선행연구를 통해 어느 정도 검증되었으나, 전통적인 작업 환경을 대상으로 한 연구에서 도출되었으므로

클라우드소싱에 모두 적용하기는 어렵다. 클라우드소싱은 상당히 큰 집단이 하나의 작업을 수행하지만, 개별 작업자들끼리는 서로 알지 못하기 때문에 집단 작업환경이면서도 동시에 극히 개인화된 작업 환경이다. 그래서 타인과의 비교가 큰 영향력을 주지 못할 것이므로 절대적 피드백이 클라우드소싱 환경에서 적합할 것으로 보인다. 아울러 여러 개의 다른 작업을 빠른 시간 동안 수행하는 클라우드소싱 환경에서는 작업자에게는 즉각적인 피드백이 지연된 피드백보다 더 유용할 것이다.

그러나 전달하는 측면이나 대상에 대해서는 그 영향력을 파악하기 어렵다. 전달 측면이나 대상은 클라우드소싱 작업자들에게 자신의 작업에 대한 판단 근거로 작용하는데, 이 근거가 어떻게 활용되는지는 쉽게 판단하기 어렵기 때문이다. 클라우드소싱에서 작업을 진행할 때 작업자는 자신의 결과물에 대해 두 가지 가능성을 갖게 된다. 의뢰자가 자신의 결과물을 거부할 경우 약속된 보상을 받지 못할 수 있지만 좋은 결과물을 제공한 경우에는 보너스를 받을 수도 있다. 그러므로 작업 결과물의 수준이 어느 정도인지를 가늠할 수 있어야 한다. 이때 긍정적인 피드백을 받았다면 두 번째 가능성에 초점을 맞추어 더욱 열심히 일할 수도 있지만, 반대로 부정적인 피드백을 받았다면 첫 번째 가능성에 초점을 두고 결과물을 더 열심히 다듬을 수도 있다. 또한 결과물의 좋고 나쁨을 평가해주는 방식과 현재 작업 결과물에

대한 정보를 알려주는 방식은 작업자로 하여금 두 가능성을 모두 고려하게 만들 수 있다. 이처럼 전달측면과 대상에 따른 서로 다른 피드백은 모두 클라우드소싱 작업자에게 유용하게 활용될 수 있지만, 이중에서 결과물의 질적 향상에 효과적인 피드백이 무엇인지는 탐구가 필요하다. 따라서 다양한 피드백 유형 중 전달 측면에 따른 긍정적 피드백과 부정적 피드백, 그리고 전달 대상에 따른 서술적 피드백과 평가적 피드백의 영향력을 비교 검증하는 것이 클라우드소싱에서 적합한 피드백 유형을 탐구하는데 적합할 것이다.

## 2.5 목표의 특성 및 효과 연구

Bandura(1989)에 따르면 목표설정은 차이를 만드는 과정(discrepancy-creating process)이기 때문에 현재를 불균형 상태로 만들고, 사람들은 균형 상태로 돌아가기 위해 목표와 현재의 차이를 줄이려고 노력하게 된다. 이처럼 목표는 작업자에게 현재 상태의 차이를 알려줌으로써 지속적으로 그 차이를 줄이려는 동기를 부여하며, 결과적으로는 목표에 가까운 성과를 만들어 낸다. Locke 와 Latham(2002)은 35년간의 목표설정이론 연구를 검토하면서 작업의 성과를 높이는 효과를 가진 요인들을 정리하여 목표설정이 성과로 이어지는 과정을 설명했다. 이들에 따르면 목표의 특성(goal core)은 성과(performance)에 영향을 주고, 그

과정에 다양한 중재변인과 목표 설정의 매커니즘이 기능한다. 먼저 목표를 설정함으로써 작업자는 지도적 기능, 동기부여 기능, 지속화 기능, 그리고 전략적 기능 등 4가지 매커니즘을 거칠 것으로 기대된다. 목표를 설정할 경우 작업자는 이러한 4가지 기능 혹은 매커니즘에 따라 작업을 완료하고자 하는 동기를 가지고 작업에 임하게 된다. 그리고 이러한 매커니즘은 목표 설정 단계로부터 작업 완료 단계로 가는 과정에서 다양한 중재변인에 따라 다르게 나타난다. 중재변인은 크게 목표 기여도(goal commitment), 피드백(feedback), 작업 복잡성(task complexity) 등 3가지가 있고, 목표 기여도는 다시 목표 중요성(goal importance)과 자기 효능감(self-efficacy)을 포함하므로 5가지로도 볼 수 있다.

이와 같이 목표설정이론은 제공한 목표의 특성이 작업성과에 영향을 미치는 과정에서 피드백을 포함하였다. 본 연구 또한 목표 특성과 피드백이 작업성과에 미치는 영향을 보고자 하지만, 목표설정이론과 같이 피드백을 중재변인으로 보기 보다는 개별적인 독립변인으로 보고자 한다. 이는 전통적인 작업환경과 클라우드소싱 작업환경의 차이 때문이다. 목표설정이론이 피드백을 중재변인으로 볼 수 있는 이유는 작업이 지속적, 반복적이고 작업자들이 하나의 목표를 지속적으로 가지는 전통적인 작업환경에 기반했기 때문이다. 일반 회사를 생각해보면 몇 개월 이상 지속되는 목표가 있고 서로 얼굴을 알고 같이 일하는 동료가

있으며, 목표를 이루기 위해 조금씩 여러 단계의 목표를 이루어 나간다. 따라서 큰 목표가 고정되어 있는 상태에서 지속적으로 피드백을 제공한다면 성과에 긍정적인 영향을 줄 수 있을 것이다. 그러나 클라우드소싱에서는 그런 환경을 만들기가 더럽다. 작업자가 참여한 작업은 이미 작은 단계로 나누어져 있고, 지속적인 목표가 있다기보다는 자신이 참여한 작업 내에서 최소한으로 요구되는 목표를 성취하려는 정도에 그치기 때문이다. 따라서 일회적인 피드백은 가능할 수 있어도 목표설정이론에서 가정하는 지속적인 피드백은 어렵다. 그리고 무엇보다 피드백이나 목표 제공이 현재 클라우드소싱에서는 흔한 인터페이스가 아니기 때문에, 피드백을 제공하는 자체만으로도 결과의 질에 영향을 줄 수 있다. 차후 연구에서는 두 변인의 관계를 탐구해볼 수 있겠으나, 현재 단계에서는 개별적인 영향력을 알아보는데 집중하고자 한다.

Latham 과 Steele(1983)의 연구 결과에 의하면, 목표 난이도가 같고 스스로 혹은 목표 설정과정에 참여하여 목표를 세운 경우에는 그렇지 않은 경우와 비교하여 유의미한 차이를 보이지 않았다. 그러나 목표가 아무런 설명 없이 단호하게 명령조로 제시된 경우에는 목표 설정 과정에 참여한 경우보다 그 성과가 하등하게 나타났다(Erez, 1986; Erez, Earley & Hulin, 1985; Erez & Kanfer, 1983). 따라서 작업 목표에 작업자의 기여도와 관여도를 높이기 위해서는 충분한 설명이 필요하다는 점을 알 수 있다. 이는 목표

특성(goal core) 중 하나인 세부성(specificity)과도 연계된다. Locke 와 Latham(1990)에 따르면 목표가 어렵더라도 세부적이고 자세한 경우에는 지속적으로 높은 성과를 냈지만, 최선을 다하라는 식의 두루뭉술한 목표는 동기부여가 되지 않고 단순히 이상적인 목표로만 보여 성과 향상에 유의미한 영향을 보이지 않았다. 그러므로 질적으로 향상된 작업 결과를 얻기 위해서는 자세하고 뚜렷한 목표를 제시해야 한다는 점을 알 수 있다. 작업자가 목표를 충분히 이해해야 목표의 중요성에 공감할 수 있고, 목표를 이룰 수 있다는 자기 효능감까지도 기대할 수 있기 때문이다.

목표의 특성은 다양하게 이해할 수 있는데, Pintrinch(2000)의 연구는 목표가 전달하는 정보에 따라 목표를 3가지 종류로 나누었다. 그에 따르면 목표는 크게 표적형 목표(target goals), 일반목표(general goals), 그리고 성취형 목표(achievement goals) 등으로 분류된다. 표적형 목표는 일반적으로 개인의 목표와 관련해 자주 연구된 유형으로, 중간고사에서 100점을 맞겠다는 등의 목표를 말한다. 표적형 목표는 개인이 자신의 결과를 평가할 수 있는 기준을 자세하게 알려주지만 왜 그 목표를 달성해야 하는지는 알려주지 못한다(Harackiewicz & Sasone, 1991). 반면 일반목표는 목표에 대한 이유를 자세하게 알려주지만 자신의 결과를 평가할 수 있는 세부적 기준을 제공해주지는 못한다. 일반목표의 예로는 인간관계나 사회적 책임감과 관련된 사회적 목표가 있다. 표적형

목표와 일반목표 사이에 성취형 목표가 있는데, 이는 개인이 성취하려는 작업을 해야 하는 이유를 설명해주면서도 자세한 기준을 제공한다. Erez의 연구들과 Locke와 Latham의 연구결과를 종합해볼 때, 필요성에 대한 설명이 있는 일반목표나 성취형 목표가 표적형 목표보다 작업자의 목표 기여도를 높일 것이라 기대할 수 있다. 그러나 클라우드소싱에서는 작업이 단기적이고 일시적이기 때문에 오히려 세부적 기준이 있는 표적형 목표나 성취형 목표가 일반목표보다 작업자의 목표 기여도를 높일 수 있을 것으로 보인다. 따라서 표적형 목표와 성취형 목표를 클라우드소싱에서 적용할 수 있는 목표 방식으로 채택하고, 그 영향력을 비교하는 것이 타당할 것이다.

아울러 단계목표(proximal goal)와 최종목표(distal goal)에 관한 연구도 진행되었다. Latham과 Seijts(1999)에 따르면 최종목표는 ‘최선을 다하라(Do your best)’ 등 규범적이거나 혹은 전체 작업을 마쳤을 때 달성해야 할 목표를 말한다. 반면 단계목표는 실제 작업 내용에 대해 자세하게 서술되어 있어 작업 중에 달성해야 하는 세부적 목표이다. 이러한 특성으로 인해 단계목표는 단계별 작업 가이드라인에 가까우며, 최종목표는 말 그대로 전체 작업이 끝난 후에 완성해야 할 목표이다. 이들의 연구 결과, 단계목표와 최종목표가 함께 제시되었을 때의 업무 성과는 최종목표만 제시되었을 때보다 유의미하게 높게 나타났다. Latham과 Seijts는

단계목표가 작업자로 하여금 일종의 내적 피드백을 유도하고, 현재 상황이 목표와 연계되어 있는지 혹은 얼마나 떨어져있는지 알려줌으로써 오류 관리(error management)를 돕는다고 설명했다. 따라서 여러 단계의 작업을 진행하여 최종목표까지 도달하는 방식의 일을 진행할 때, 큰 목표만 주는 것보다는 단계별로 자세하고 즉각적인 정보를 얻을 수 있는 단계목표를 제공하는 것이 효과적일 수 있다. 이는 Locke 와 Latham(1990)의 연구 결과와도 연결되는데, 이들은 목표가 세부적일 때 지속적으로 높은 성과를 낼 수 있다고 밝혔다. 단계목표와 최종목표를 함께 제시하는 방식은 목표를 다각적이고 세부적으로 제공하게 되므로 Locke 와 Latham 이 주장한 바와 같이 목표의 세부성이 결과물의 질을 높이는데 기여한다고 볼 수 있다. 클라우드소싱은 비대면 작업환경으로 작업자가 작업 행위와 관련된 정보를 얻는데 한계가 있으므로, 세부적인 목표가 클라우드소싱에서도 유용할 것으로 보인다.

이와 같이 목표와 관련된 연구들을 살펴보면 목표는 전달하는 정보 및 달성 시기에 따라 유형을 구분할 수 있었다. 그 결과는 아래 표와 같다.



표 2. 목표의 유형

기준	유형	속성
정보유형	표적형 목표	명확한 결과물 평가 기준만 존재함
	일반형 목표	작업을 달성해야 하는 이유만 설명함
	성취형 목표	명확한 결과물 평가 기준과 이유를 설명함
달성시기	단계목표	작업 과정에서 달성해야 할 목표
	최종목표	작업 완료 시에 달성해야 할 목표

목표설정이론은 설정된 목표가 다양한 중재변인을 통해 성과로 이어지는 과정을 설명했으며, 이러한 과정에 있어 목표 자체의 특성 및 목표와 중재변인과의 관계가 중요하다고 주장했다. 현재 클라우드소싱 환경은 작업자에게 해결해야 하는 작업만을 제공하거나, 작업에 대해 간단히 소개하는 작업 설명(instruction)이 많아 금전적 보상 외에 뚜렷한 목표를 주지는 못하고 있다. 금전적 보상의 효과는 앞서 살펴본 것처럼 영향력이 제한되어 있어 새로운 목표가 필요하다. 따라서 본 연구에서는 각각의 목표 속성이 클라우드소싱 작업 환경에서 어떠한 결과를 만들어내고, 그 차이는 어떠한지 살펴보고자 한다.

### 3. 예비연구

#### 3.1 실험목적

본 연구에 앞서 클라우드소싱 작업에 대한 커뮤니케이션이 작업자들의 작업 행위에 영향을 미치는지 알아보기 위하여 먼저 예비연구(pilot study)를 실시하였다. 클라우드소싱 환경에서는 일반적으로 작업자와 직접적으로 커뮤니케이션하는 경우가 드물어 작업자와의 커뮤니케이션이 실제로 어떠한 결과를 가져오는지 쉽게 짐작하기 어렵다. 따라서 본 실험에 들어가기에 앞서 클라우드소싱 환경 내에서 작업자와의 커뮤니케이션이 어떤 효과를 가지는지 살펴볼 필요가 있었다. 그래서 예비연구에서는 구체적으로 형식과 내용이 정해진 메시지를 제공하기 보다는 커뮤니케이션 가능성만을 언급하는 방식으로 실험을 진행하였다. 구체적인 메시지를 제공할 경우 그 내용에 따른 영향력이 나타날 수 있다. 그러므로 실제 커뮤니케이션의 전단계로서 커뮤니케이션 가능성을 언급하고 이에 따른 작업자의 행동 변화를 탐색하여 커뮤니케이션이 갖는 효과를 간접적으로 살펴보고자 했다.

표 3. 예비연구 실험 조건

	어려운 작업(Complex)	쉬운 작업(Simple)
검토조건 (Evaluation)	EC(검토 x 어려운 작업)	ES(검토 x 쉬운 작업)
비검토조건 (Independent)	IC(비검토 x 어려운 작업)	IS(비검토 x 쉬운 작업)

예비연구에서는 실제로 작업 결과물에 대한 특정 메시지를 전달하는 것이 아니라, 검토할 것이라는 사실만을 알려주고 어떤 행동을 보이는지를 살펴보았다. 이에 따라 검토 가능성을 언급한 검토 조건과 그렇지 않은 비검토 조건을 나누었다. 또 커뮤니케이션의 영향력이 작업의 난이도에 따라 다르게 나타나는지도 살펴볼 필요가 있었다. 쉽고 단순한 작업을 할 때 누군가가 검토를 한다는 사실이 작업에 집중하도록 할 수도 있지만, 단순작업에 검토까지 한다는 것이 지나친 간섭으로 느껴질 수도 있기 때문이다. 또 어렵고 복잡한 작업을 할 때도 제3자의 검토 사실은 작업을 포기하게 하지 않는 역할을 할 수도 있지만, 작업 난이도가 높다보니 별다른 영향력이 없을 수도 있다. 그러므로 검토 조건과 비검토 조건, 그리고 쉬운 작업과 어려운 작업을 결합하여 2x2 조건을 설정하고 각 조건에서 작업자의 행동 변화를 관찰하였다. 예비연구의 실험내용은 클라우드소싱에서 자주 활용되는 분류 작업이었으며, 작업자의 행동 변화를 관찰하기 위한 척도로 작업시간과 정확도를 선정하였다. 분류작업에서 중요한 점은 빠른 시간 안에 정확하게 대상을 분류해내는 것이기 때문에, 이 두 가지 척도를 통해 작업자의 행동이 바람직한 방향으로 움직였는지 파악하고자 했다. 정확도는 전체 답안 중에서 정답을 정답이라고 답한 경우와 오답을 오답이라고 답한 경우의 비율을 측정하였다.

## 3.2 실험내용

예비연구의 실험내용은 그림 분류작업으로 클라우드소싱에서 자주 사용되는 작업 중 하나이다. 예비연구 참여자들은 12장의 남성 정장 사진을 보고, 제시된 조건에 따라 이를 분류하는 작업을 수행했다. 커뮤니케이션 가능성을 언급하는 검토조건에서는 작업자가 작업을 시작하기 전 읽게 되는 가이드 화면에 해당 메시지를 노출하였다. 작업 결과물에 대해 판단할 제3자로 구성된 평가 집단(evaluator group)이 10점 만점으로 결과물에 대해 점수를 매길 것이며, 평가 점수가 7점 이상일 때는 보너스가 지급될 수 있다는 것이 메시지의 내용이었다. 이 메시지는 작업 결과물에 대한 미래 커뮤니케이션의 가능성을 언급하는 동시에 그 자체가 의뢰자와 작업자 간 커뮤니케이션이기도 하다. 반면 비검토조건에서는 최종 제출 전 검토해달라는 일반적인 안내 문구만 제시하여 검토조건의 메시지와 구별하였다.

아울러 작업의 난이도를 달리 하기 위해 같은 사진에 대해 다른 조건을 제시하였다. 쉬운 작업 조건에 해당하는 경우 12장의 사진 중 단추가 2개인 정장을 분류하도록 했다. 반대로 어려운 작업 조건에 해당하는 경우 미국 스타일 정장에 대한 설명 및 체크리스트를 제공하고, 이에 따라 해당 정장을 분류하도록 했다. 쉬운 작업 조건에서 작업자는 하나의 사진에서 단추의 개수라는 한 가지 정보만 판단하지만, 어려운 작업 조건에서 작업자는 3가지의

정보를 판단해야 한다. 또한 그 3가지 정보가 단추 개수와 같이 명확하고 단순한 정보가 아니기 때문에 상대적으로 많은 인지적 노력이 필요하다. 이처럼 두 작업이 작업자에게 요구하는 인지적 자원의 수준이 다르기 때문에 두 작업의 난이도 간에는 충분한 차등이 있다고 판단되었다. 위와 같이 검토 메시지와 난이도를 종합하여 아래 그림과 같이 작업자에게 작업지시문을 제공하였고, 그 대상은 <그림 3>에 나타난 12장의 남성 정장 사진이었다.

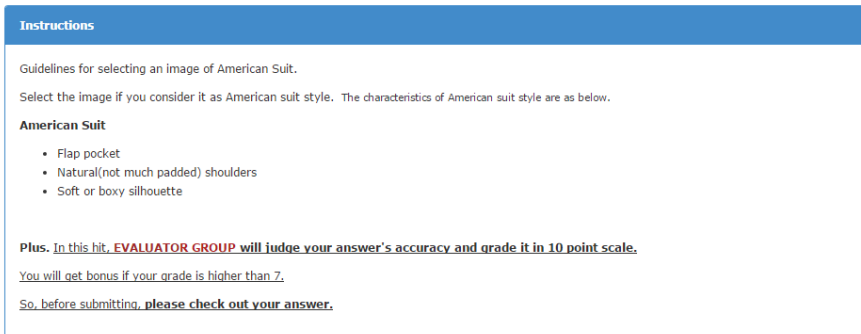


그림 1. 어려운 작업 x 검토 조건의 작업 지시 화면



그림 2. 쉬운 작업 x 비검토 조건의 작업 지시 화면



그림 3. 분류작업 사진

### 3.3 실험절차

예비연구는 아마존 메커니컬 터크(Amazon's Mechanical Turk)에서 2014년 7월 15일부터 7월 21일까지 진행되었다. 해당

사이트에서 진행되는 작업 한 단위를 HIT(Human Intelligence Task)라고 부르는데, 실험 참여자를 모집한 HIT 의 제목은 ‘Suit Style\_American Suit’였다. 실험 진행 전 수집 데이터가 일정한 수준 이상이 되도록 실험 참여자 모집 시 2가지 조건을 제시하였다. 먼저 다른 HIT 에서 해당 참여자가 제공한 결과물이 채택된 비율이 98% 이상이어야 하고, 그러한 HIT 가 5000개 이상이어야 예비연구에 참가할 수 있었다. 실험에 참여하여 답안을 제출하고 나면 연구자가 보수를 지급하고 실험을 완료하였다.

### 3.4 연구가설

위와 같은 목적과 내용, 절차를 가진 예비연구에서 알아보고자 하는 연구가설은 아래 4가지와 같다.

- H1. 검토 조건에서의 작업시간이 비검토 조건에서의 작업시간보다 짧을 것이다.
- H2. 쉬운 작업에서의 작업시간이 어려운 작업에서의 작업시간보다 짧을 것이다.
- H3. 검토 조건에서의 정확도는 비검토 조건에서의 정확도보다 높을 것이다.
- H4. 쉬운 작업에서의 정확도는 어려운 작업에서의 정확도보다 높을 것이다.

### 3.5 실험결과

위와 같이 4가지 조건에서 각각 30명을 모집하여 실험을 진행한 결과 120건의 데이터를 확보하였고, 그 중 불성실한 답변 및 이상치를 제외하고 100건의 데이터를 최종 분석에 사용하였다. 집단간 비교를 위하여 t-검정을 실시한 결과, 본인의 작업 결과물이 검토될 것이라는 메시지를 받은 작업자들은 그렇지 않은 경우보다 더 많은 시간을 투자하였다. 검토 조건의 평균 작업 시간( $M=112.50$ )은 비검토 조건( $M=46.62$ )보다 길어( $t(53.145)=4.41, p < .05$ ), 검토 가능성은 작업 시간에 영향을 주고 있었다. 그러나 검토 가능성이 작업 정확도를 높이지는 못했다. 하지만 검토 조건의 평균 작업 정확도( $M=0.43$ )는 비검토 조건의 평균 작업 정확도( $M=0.472$ )보다 오히려 낮았다( $t(97.06)=-2.03, p < .05$ ). 어려운 작업 조건의 작업시간( $M=71.26$ )과 쉬운 작업의 작업시간( $M=59.81$ )간에 유의한 차이가 없어, 작업시간에 난이도가 영향을 주지는 않고 있었다. 또한 작업 정확도에서도 어려운 작업의 정확도( $M=0.467$ )와 쉬운 작업의 정확도( $M=0.443$ ) 간 차이가 없었다.



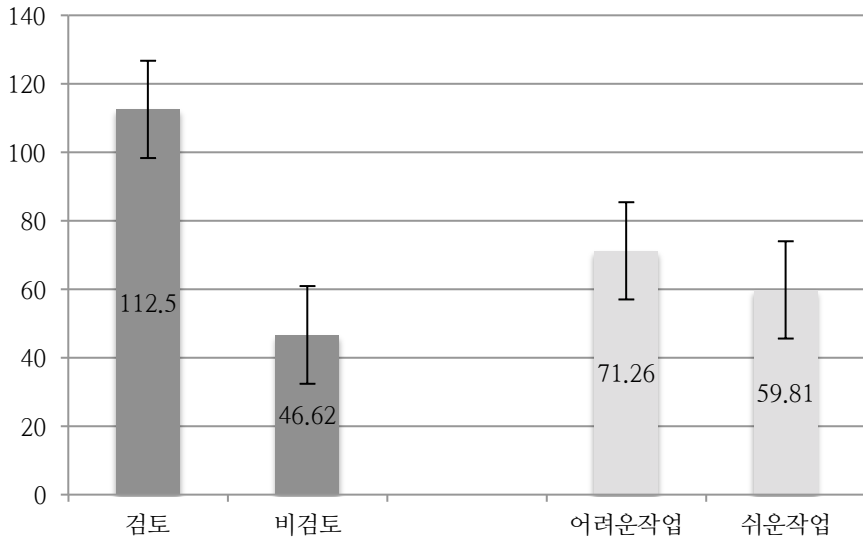


그림 4. 예비연구\_그룹별 작업시간 비교

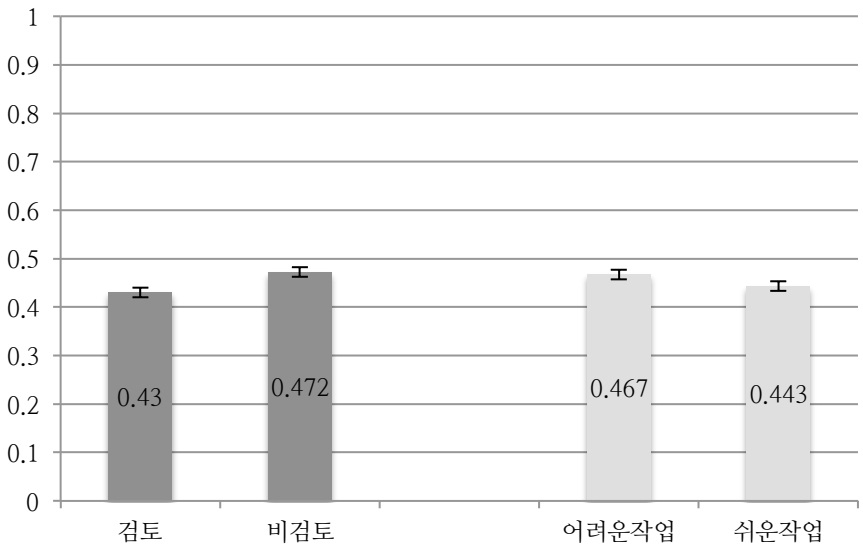


그림 5. 예비연구\_그룹별 작업 정확도 비교

위와 같은 실험 결과 클라우드소싱 내 작업자들은 자신의 결과물이 검토될 것이라는 가능성만으로도 더 많은 시간을 투자하는 모습을 보여, 작업과 관련된 커뮤니케이션이 작업 행동에 영향을 준다는 점을 알 수 있었다. 그러나 커뮤니케이션이 정확도를 높이지는 못하고 있었다. 검토 조건보다 비검토 조건의 정확도가 더 높았고, 무엇보다 조건별 정확도가 모두 0.5 이하로 나타나 50% 이하의 정확도를 보였다. 이러한 결과는 커뮤니케이션의 활용에 대해 흥미로운 시사점을 던져준다. 커뮤니케이션은 작업자의 행동을 바꿀 수 있지만, 그 행동 변화가 결과물의 향상으로 이어지기 위해서는 작업자와의 커뮤니케이션 자체보다 그 안에 담긴 내용에 주목해야 한다는 것이다. 제3자가 작업에 대해 지켜보고 소통한다는 점에서 더 나아가 구체적으로 어떻게 소통할 것인지를 고민할 때, 커뮤니케이션이 결과물의 질을 높일 수 있을 것이다. 본 연구는 선행연구 검토에 기초하여 구체적인 커뮤니케이션 방안 중 피드백과 목표에 주목하였으며, 여러 방안의 비교 분석을 통해 클라우드소싱에 적합한 커뮤니케이션 방안을 알아보고자 한다.

## 4. 연구문제 및 가설

선행연구 검토에서 살펴본 기존 클라우드소싱 품질 관리 방식은 작업 자체에 초점을 맞추고 있어 작업 결과물의 질적 향상에

한계를 가지고 있었다. 이러한 한계를 극복하기 위해 본 연구는 작업자와의 커뮤니케이션이 갖는 가능성에 주목하고, 선행연구들에서 작업자와의 가능한 커뮤니케이션을 피드백과 목표로 유형화하였다. 그리고 예비연구에서 작업자와의 커뮤니케이션이 작업자의 행동 변화에 어떤 영향을 주는지 알아보았다. 그 결과 작업자들은 결과물이 검토된다는 메시지를 받았을 때 작업에 더 많은 시간을 투자하며 행동을 바꾸고 있었다. 하지만 작업 정확도는 차이를 보이지 않아 더 구체적인 커뮤니케이션이 필요함을 알 수 있었다. 따라서 본 연구에서는 예비연구를 확장하여 다양한 커뮤니케이션 방식이 결과물의 질적 향상에 미치는 영향력을 비교하고자 한다. 연구문제 1에서는 다양한 피드백 방식이 결과물의 질을 높이는 정도를 비교하고, 연구문제 2에서는 여러가지 목표 제시가 결과의 질적 향상에 미치는 영향력을 비교한다.

연구문제 1. 어떠한 형태의 피드백이 결과의 질을 높이는가?

가설 1.1 부정적 피드백 조건에서 결과물의 질은 긍정적 피드백 조건에서보다 높을 것이다.

가설 1.2 서술적 피드백 조건에서 결과물의 질은 평가적 피드백 조건에서보다 높을 것이다.

한주희와 고수일(2005)의 연구에 따르면 피드백의 긍정적 혹은 부정적 내용이 결과물에는 큰 차이를 보이지 않았다고 했다. 그러나 이들의 연구는 팀 기준의 피드백이었기 때문에 긍정적 혹은 부정적 피드백이 전반적인 영향을 주지 않았을 수 있다. 크라우드소싱은 여러 명이 모여 있으나 서로 보지 못한 채 혼자 작업하는 방식이기 때문에 내적 동기에 영향을 줄 수 있는 피드백을 고려할 필요가 있다. 아울러 반사실적 사고방식 연구는 부정적 상황이 반사실적 사고방식을 불러일으키고, 이는 행위의도를 통해 실제 행동으로 이어진다고 설명하고 있다. 이런 점을 생각해보다면 같은 내용을 부정적으로 표현하여 부정적 상황을 인지하도록 자극하는 방식이 오히려 칭찬과 같은 긍정적 피드백보다 성과에 정적 영향을 끼칠 수도 있다. 이와 같은 이유로 가설 1.1을 통해 긍정적 피드백과 부정적 피드백의 영향력을 검증해보고자 한다.

아울러 과정에 대한 피드백과 결과물에 대한 피드백을 살펴볼 필요가 있다. 한주희와 고수일(2005)은 과정 중심의 피드백이 내적 동기를 유지하는데 효과적이라고 보고했으며, 권은주와 김성일(2003)은 수행의 질을 높이기 위해서 작업자 본인 혹은 과제 자체에 대한 피드백이 필요하다고 주장했다. Harackiewicz 등(1987)도 과정중심 피드백이 중요하며 효과적이라고 말한 바 있다. 이와 관련해 서술적 피드백은 작업과정에 대해 서술하고

작업에 필요한 정보를 제공하므로 과정에 집중한다면, 평가적 피드백은 결과물의 정답 여부에 집중한다. 선행연구와 마찬가지로 서술적 피드백의 효과가 평가적 피드백보다 높을 경우 클라우드소싱에서 작업과정 데이터를 통해 지속적이고 즉각적으로 피드백을 주는 시스템을 고안해낼 필요가 있을 것이다. 따라서 가설 1.2를 통해 과정 중심 대 결과 중심 피드백의 영향력을 알아보고자 한다.

연구문제 2. 어떠한 형태의 목표가 결과의 질을 높이는가?

가설 2.1 단계목표와 최종목표를 함께 제시하는 조건에서 결과물의 질은 최종목표만 제시하는 조건에서보다 높을 것이다.

가설 2.2 성취형 목표를 제시한 조건에서 결과물의 질은 표적형 목표가 제시된 조건에서보다 높을 것이다.

목표의 특성이 성과에 미치는 영향 또한 본 연구의 관심사이다. 클라우드소싱의 두 주체인 의뢰자(requester)와 작업자가 가진 기본적 목표는 다른데, 목표를 제공하는 방식이나 형태를 통해 두 목표 간 간극을 줄여보고자 한다. 클라우드소싱에 과제를 올리는 의뢰자는 빠르게 데이터를 모으려는 목표가 있고, 클라우드소싱에서 일하는 사람인 작업자는 금전적 보상 등 개인적

목표를 가지고 있다. 이러한 차이는 일반 회사와 같은 집단 작업 환경에서도 비슷하게 존재하지만, 집단의 구성원들은 지속적인 관계를 유지해야 하고 집단의 목표 달성 여부가 개인에게 영향을 주기 때문에 두 목표가 어느 정도는 수렴할 수 있다. 반면 클라우드소싱의 작업자와 의뢰자는 일시적인 관계이고 의뢰자의 목표가 달성되지 않아도 작업자가 받을 영향력은 보너스를 받지 못하는 정도에 그치기 때문에 두 목표를 수렴하기가 상대적으로 더 어렵다. 따라서 추가적인 목표를 제공하고, 그 속성을 달리 하여 어떤 목표가 작업의 질을 높이는지 알아보고자 한다.

Latham 과 Seijts(1999)의 연구에 따르면 단계목표, 즉 작업자가 지금 하고 있는 작업 과정에서 성취해야 할 작은 목표들을 최종목표와 함께 제시하는 경우 최종목표만 제시했을 때보다 작업성고가 향상되었다. 이는 단계목표가 작업자의 오류 관리를 돕기 때문으로 보이는데, 본 연구도 이들의 연구를 차용하여 클라우드소싱에서도 같은 효과가 나타나는지 알아보고자 한다. 이러한 방식은 단계별로 자동화된 결과물 관리 시스템을 만들 때도 유용하게 쓰일 수 있을 것으로 기대된다. 작업 소개 단계에서만 최종목표를 언급하는 기존 시스템에 비해 작업자가 완료한 작업 결과를 토대로 단계목표를 선택하여 제공하는 방식을 예상해볼 수도 있다. 이와 관련된 가설은 2.1과 같다.

아울러 가설 2.2와 같이 성취형 목표가 표적형 목표보다 결과의 질에 긍정적인 영향을 끼칠 것이라고 예상한다. Erez 와 Kanfer(1983) 및 Erez 등(1985) 등의 연구에서 목표가 아무런 설명 없이 단호하게 명령조로 제시된 경우에는 그렇지 않은 경우보다 성과가 떨어졌다. 또한 Locke 와 Latham(1990)에 따르면 최선을 다하라(Do-your-best) 식의 두루뭇술한 목표보다는 세부적 목표가 성과 향상에 효과적이었다. 이러한 연구 결과들을 볼 때, 목표를 왜 추구해야 하는지에 대한 명확한 설명과 함께 현재 자신이 작업을 제대로 수행하고 있는지에 대한 기준을 제공할 때 성과에 긍정적인 영향을 미친다고 볼 수 있다. 성취형, 표적형, 일반형 목표는 이러한 목표를 추구해야 하는 이유와 기준 제공에 따라 분류되어 있다. 성취형의 경우 이유와 기준이 모두 있는 유형의 목표이지만, 표적형은 기준만 제시되어 있고 일반형은 이유만 제시하고 있다. 따라서 이유와 기준을 모두 제시한 성취형 목표가 다른 두 목표보다 결과물의 질적 향상에 더 큰 영향력을 끼칠 것이라 예상할 수 있다. 그러나 여기에서는 성취형 목표와 표적형 목표만 비교하기로 한다. 클라우드소싱은 익명적인 작업환경이고 개별 작업자를 통제하기 어려운 구조이며, 무엇보다 작업자들은 금전적 보수를 포함한 개인적인 목적을 가지고 작업에 참여한다. 따라서 작업을 해야 하는 이유만을 별도로 제공하는 것보다는 자신의 작업에 대한 구체적인 목표를 제시하는 것이 더 적합하다고

판단된다. 단 앞선 연구에 근거할 때 왜 작업을 해야 하는지 이유를 알 경우 목표 제공 효과가 커질 수 있으므로, 성취형 목표와 표적형 목표를 비교하고자 한다.

## 5. 연구방법

### 5.1 참여자 모집

본 연구의 실험의 참여자는 아마존 메커니컬 터크에서 모집하였다. 선행연구 검토에서 설명한 것과 같이 메커니컬 터크는 대표적인 클라우드소싱 사이트로, 사용자가 많고 작업을 디자인하기 용이해 클라우드소싱 연구에서 많이 사용된다. 작업 시작 전 작업자의 능력을 검증하는 테스트를 설정하거나 해당 작업자의 이전 기록을 검토하여 자동적으로 필터링할 수 있기 때문에 데이터의 질적 수준을 관리할 수 있다. 또한 이미지 라벨링부터 설문조사까지 다양한 작업을 할 수 있도록 기본적인 인터페이스를 제공하기 때문에 실험을 진행하기 용이하다. 이러한 이유로 메커니컬 터크에서 실험 참여자를 모집하여, 2015년 5월 26일부터 2015년 6월 10일까지 실험을 진행했다.

연구문제에서 밝힌 대로 본 연구는 피드백과 목표를 각각 4개 유형으로 지정하되 2개씩 짝지어 비교했다. 집단 간 비교가 용이하도록 각 유형은 20명 내외로 하며, 비교집단 간 참여자 수를



유사하게 하였다. 아울러 보수의 양이 미치는 영향력을 함께 살펴보기 위해, 본 연구는 2가지 보수 지급 조건에서 동일한 실험을 진행했다. 본 연구의 실험 내용은 블로그 혹은 논문에서 일부 발췌한 텍스트를 읽고 길이를 줄이는 작업으로, 일반적인 클라우드소싱 작업에 비해 상당한 노력을 요구하므로 지나치게 적은 보수는 작업의 질에 악영향을 줄 수 있다. 따라서 보수의 기본 수준을 일반적으로 메커니컬 터크에서 지급하는 것보다 높게 설정했다. Bernstein 등도 응답률 등을 고려하여 기존보다 넉넉하게 배당하여 탐색 단계에서는 줄어든 글자 당 \$0.08, 교정 단계는 \$0.05, 확인 단계는 \$0.04를 지급하였다. 이와 같은 선례를 참고하되, 본 연구는 실험의 편의를 위해 줄어든 비율(%) 당 지급하는 방식을 선택하였다. 1000자의 1%인 10자를 줄일 경우에만 글자당 \$0.03를 지급하면 \$0.3를 지급하는데, 본 실험은 1%당 \$0.3을 지급하므로 방식은 다르나 실질적으로 제공하는 보수의 양은 유사하다.

## 5.2 실험내용

예비연구에서 나타난 것과 같이 클라우드소싱 작업자들이 익숙하지 않은 종류의 작업을 이용할 경우 실험 결과가 왜곡될 수 있다. 작업자들이 피드백이나 목표 등 실험이 측정하고자 하는 변인에 영향을 받아야 하는데, 검증되지 않은 작업을 사용할 경우

내용이 낮설거나 지나치게 어려워 결과에 부정적인 영향을 줄 수 있기 때문이다. 이러한 점을 고려하여 선행연구에서 살펴보았던 Bernstein 등(2010)의 연구의 실험내용을 차용하였다. 이들은 크라우드소싱을 이용해 텍스트 교정 작업을 실시했는데, 텍스트 단축(shortn), 교정(crowdproof), 그리고 복합적 교정 작업(human macro) 등 3가지 작업이 이루어졌다. 이 중에서 본 연구는 텍스트 단축 작업을 차용했다. 해당 작업은 앞선 연구에서 그 유용성이 입증되었고, 결과물의 길이가 지나치게 길거나 짧아진 부분을 알려주는 등 기준이 명확해 피드백이나 목표를 제공하기 용이하기 때문이다.

작업자에게 제공된 텍스트는 Bernstein 등의 연구에서 사용된 텍스트를 사용했다. 텍스트의 내용과 수준은 매우 다양해 실험 처치가 적절히 나타나는 텍스트가 필요한데, 단어나 문장 구조가 어렵거나 텍스트의 주제가 낮설다면 텍스트의 내용 때문에 작업이 더뎠고 결과물의 길이 상대적으로 낮아질 수 있고, 연구 목적에 따른 실험을 진행하기 어려울 수 있기 때문이다. 반면 지나치게 쉬운 텍스트라면 피드백이나 목표의 제공 없이도 충분히 좋은 결과물을 만들 수 있어 그 영향력이 잘 측정되지 않을 수도 있다. 그러므로 선행연구에서 검증된 텍스트를 사용하는 것이 적절하다고 판단했다. 텍스트 난이도를 상쇄할 수 있도록 작업자에게 난이도가 높은 텍스트와 낮은 텍스트를 무작위로 부여하고자 했는데,

난이도가 높은 텍스트로는 Ullmer 와 Ishii(1997)가 UIST 에  
게재했던 metaDESK 관련 논문에서 발췌한 부분을 선택했다. 또  
난이도가 낮은 텍스트로는 IT 관련 전문  
블로그(techcrunch.com)에서 iPad 에 대한 포스팅을 사용했다.  
아울러 메커니컬 터크의 인구통계학적 특성을 고려하여 영어  
텍스트를 사용했다. Ross 등(2010)에 따르면 2009년을 기준으로  
메커니컬 터크 작업자들의 국적 비중은 미국, 인도, 캐나다, 영국,  
필리핀 순으로 나타나 작업자 대부분이 영어권 거주자이다. 이와  
같은 메커니컬 터크의 인구통계학적 특성을 고려해보면 영어  
텍스트가 적절하다고 판단했다.

### 5.3 실험절차

본 연구는 Bernstein 등의 ‘탐색-수정-확인’ 패턴을 차용하되,  
탐색 단계는 생략하고 수정 단계에 집중한다. Bernstein 등의  
연구는 3단계 작업흐름 자체의 효과에 대해 검증한다면, 본 연구는  
그 흐름 안에서 작업자와의 커뮤니케이션의 효과에 주목하기  
때문이다. 따라서 미리 지정된 짧은 단락을 제공하는 수정 단계를  
중심으로 실험을 진행하고, 수정 단계를 수행한 사람과 확인  
단계를 수행한 사람이 다른 Bernstein 등의 연구와 달리 한  
작업자가 스스로의 작업을 검토하는 방식으로 진행되었다.

본 연구는 2단계 구조를 통해 작업자가 스스로 자신의 답안을 평가하도록 유도한다. 즉, 메시지를 읽고 본인이 1단계에 제출한 답안을 수정할 기회를 제공하기 때문에, 수정 단계와 확인 단계가 통합되어 있다고 볼 수 있다. 만약 메시지를 읽고도 별다른 차이가 나타나지 않는다면 해당 조건에서의 이루어진 작업자와의 커뮤니케이션은 결과물의 질적 향상에 효과적이지 않다는 결론이 가능하다. 실험은 메커니컬 터크에서 하나의 HIT 로 등록하여 진행했다. 여기서 HIT 란 Human Intelligence Task 의 약자로 메커니컬 터크에 등록하는 작업 하나의 단위를 뜻한다. 본 실험에서는 보수조건 \$0.15와 \$0.3의 실험 2 가지를 진행하므로, 총 2개의 HIT 가 메커니컬 터크에 등록되었다.

Text Shortening
figure1690@snu.ac.kr

### Guideline

- You'll go through steps below and in total **2 answers are required** to complete this HIT.

#	Step
1	In your 1st step, a paragraph will be shown.
2	Shorten the paragraph as much as you can, and write it in your answer box below.
3	Shortly after submitting the 1st answer, a message about your answer will appear.
4	When you come back to working page again, you are asked to revise your prior answer.
5	Submitting the 2nd answer, you'll get completion code.
6	Put the code in MTurk and get paid!

### Your payment

is based on the percentage of how much you've shortened the text.  
 ex) If you shorten the text into 90% of original text(shortened text is 10% of original text), your payment is 10 \* reward per assignment.

Do you understand and want to proceed?

Yes
No

그림 6. 작업 가이드라인

Text Shortening
figure1690@snu.ac.kr

**Check your GOAL before starting work.**

Check out whether you removed ...

a) Overlapped words b) Unnecessary sentences

**Do you understand your goal clearly?**

Yes

그림 7. 목표 조건에서의 사전 목표 화면

Text Shortening
figure1690@snu.ac.kr

**1st step**

Please shorten the paragraph below

The graphical user interface (GUI) has proven both a successful and durable model for human-computer interaction which has dominated the last decade of interface design. At the same time, the GUI approach falls short in many respects, particularly in embracing the rich interface modalities between people and the physical environments they inhabit. 6 Systems exploring augmented reality and ubiquitous computing have begun to address this challenge. However, these efforts have often taken the form of exporting the GUI paradigm to more world-situated devices, falling short of much of the richness of physical-space interaction they seek to augment. In this paper, we present research developing "Tangible User Interfaces" (TUIs) - user interfaces employing physical objects, instruments, surfaces, and spaces as physical interfaces to digital information. In particular, we present the metaDESK system (Figure 1), a graphically intensive system driven by interaction with graspable physical objects. In addition, we introduce a prototype application driving an interaction with geographical space, TangibleGeospace, to illustrate our approach. The metaDESK effort is part of the larger Tangible Bits project [8]. The Tangible Bits vision paper introduced the metaDESK along with two companion platforms, the transBOARD and ambientROOM. Together, these platforms explore both graspable physical objects and ambient environmental displays as means for seamlessly coupling people, digital information, and the physical environment.

your answer here

Time left  
09 39  
MINUTESSECONDS

그림 8. 텍스트 단축 작업 화면

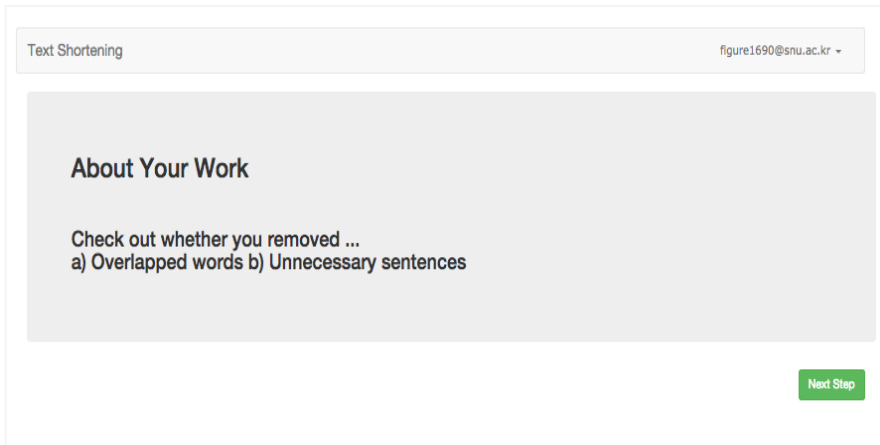


그림 9. 커뮤니케이션 메시지 예시\_표적형 목표

참여자들은 메커니컬 터크에 등록된 본 연구의 실험용 웹사이트에 접속하여 사용자로 등록했다. 이때 참여자를 구분하고, 실험 데이터 수집에 대해 동의를 구하며 실험 디브리핑(Debriefing)을 진행하기 위하여 이메일 주소를 수집하였다. 참여자는 사용자로 등록할 때 사이트로부터 실험조건과 텍스트를 하나씩 무작위로 부여받는데, 본인에게 할당된 조건은 알 수 없다. 본 실험은 비대면 참여자를 대상으로 하기 때문에 실험 참여에 대한 동의는 해당 사이트에 접속하여 사용자 등록을 하는 과정에서 간접적으로 이루어지지만, 작업 절차와 내용을 설명한 후 이러한 작업에 참여하겠다는 동의를 사이트 내에서 한번 더 받았다. 동의를 거친 후에 참여자는 본격 실험에 참여하게 되었다.

이후에는 개인에게 부여된 조건, 즉 피드백과 목표 조건에 따라 실험이 진행되었다. 실험의 기본 절차는 두 조건에서 같지만, 피드백과 달리 목표는 참여자가 미리 숙지해야 하기 때문에 <그림 7>과 같이 1번째 텍스트 단축 작업 전에 목표를 숙지하는 단계를 거쳤다. 이후 <그림 8>과 같이 참여자는 1단계에서 주어진 텍스트를 읽고 단축 작업을 한 다음 제출하는데, 이후 15초 동안 작업 결과를 처리 중이라는 메시지에 노출되었다. 실제로는 저장된 메시지 중 참여자의 조건에 해당되는 메시지가 제공되는 구조이지만, 참여자들이 실제로 자신의 답변에 대한 피드백 혹은 목표를 제공받겠다고 생각하도록 이러한 장치를 구성하였다.

이후 참여자는 스스로 해당된 조건에 따라 <그림 9>와 같이 피드백 혹은 목표 메시지를 받게 되었다. 충분히 메시지를 숙지한 후 다음 단계로 넘어갈 수 있도록 하기 위해 메시지를 다 읽은 후 다음 페이지로 넘어가는 버튼을 눌러 2단계로 넘어가도록 하였다. 이때 단축해야 하는 텍스트는 동일하나 1단계에서 제출했던 답안을 함께 제시하여 참고하도록 하였다. 2단계에서 실험 참여자는 이전 페이지에서 받은 메시지를 고려하여 1단계 답안의 내용을 수정하고 이를 제출하도록 요구받았다.

2단계 답안을 제출한 후에는 종료화면으로 이동하는데, 아마존 메커니컬 터크에 입력할 완료 코드를 부여받았다. 완료코드는 메커니컬 터크에서 외부 링크를 통해 작업을 진행할 때 작업자가

실제로 그 작업을 했는지를 판별하기 위해 사용하는 숫자코드로 연구자가 임의로 설정했다. 참여자가 이를 복사하여 메커니컬 터크에 연구자가 게시한 HIT 내 입력란에 입력하면, 연구자가 완료코드가 정확한지와 해당 작업자가 성실하게 작업을 수행했는지 확인했다.

작업자가 완료코드를 입력하면 HIT 관리 페이지에는 작업자의 일련번호 및 완료코드가 들어간 데이터 테이블이 제공되었다. 연구자는 24시간마다 이 테이블을 확인하고, 중복 작업자가 있는지 점검한 후 보수를 지급했다. 동시에 연구자는 아래 흐름에 따라 최종 제출된 작업 결과물 및 작업시간, 비용 등의 데이터를 함께 저장하였다. 또한 저장된 결과물들을 검토한 후 부적합하다고 판단된 결과물을 제외한 데이터를 데이터베이스에 저장하였다.

## 5.4 실험처치

앞서 설명한 흐름대로 실험을 진행하면서 작업자는 자신이 배당된 조건의 피드백이나 목표를 제공받았다. 피드백과 목표의 유형에 따라 구체적으로 제공되는 메시지는 아래 <표 4> 및 <표 5>와 같다. 동일한 결과물에 대해 긍정적 피드백은 사용자가 지금까지 얼마나 잘 해왔는지에 집중하지만, 부정적 피드백은 아직 부족한 부분에 초점을 두었다. 또한 서술적 피드백은 제출한 작업 결과 및 작업 과정에 대한 정보를 제공하므로, 1차 제출 결과물의



단어 수를 수집해 원래 텍스트에서 얼마나 줄어들었는지에 대한 정보를 서술하였다. 반면 평가적 피드백은 현재 결과물이 정답인지 아닌지만 판단하여 알려주었다. 이 때 결과물의 기준보다 짧거나 길 수 있음을 고려해 1차 제출 결과물의 단어 수에 따라 두 메시지 중 하나를 제공하였다. 목표와 관련해 단계목표 및 최종목표 조건에서는 작은 목표들의 리스트를 제공하는 동시에 원래 의미를 유지하는 선에서 텍스트를 줄여야 하고 최선을 다해야 한다는 최종목표를 전달하며, 최종목표 조건에서는 작은 목표들 없이 최종 목표와 관련한 메시지만 전달했다. 표적형 목표는 작업자가 스스로의 작업을 평가할 수 있는 기준을 제공하지만 성취형 목표는 표적형 목표에 더해 작업의 이유를 설명하는 메시지를 제공했다.

표 4. 피드백 조건별 제공 메시지

조건	실제 제공 메시지
긍정적 피드백	Looks like you've worked hard to save the original meaning. But please look into it in the next page, and find out if there's anything to be shortened before final submit.
	Looks like you've worked hard and shortned the text more than we expected. But please look into it in the next page, and find out whether the original meaning was not damaged.
부정적 피드백	Looks like you need to work more to shortend the text to the length we wanted. Please look into it in the next page, and check if there's anything to be shortened before final submit.

	Looks like you need to work more to keep the original meaning of the paragraph. Please look into it in the next page, and check whether the original meaning was not damaged.
서술적 피드백	You've shortened the text into 340 characters.
평가적 피드백	Your answer was longer than we expected.  Your answer was shorter than we expected.

\* 해당 메시지는 실제 작업 결과물의 글자 수를 측정하여 제공하였다.

## 표 5. 목표 조건별 제공 메시지

목표 조건	실제 제공 메시지
단계목표	a) Find out additional adjectives and sentences b) Pick unnecessary wordy phrase in paragraphs c) Remove those unless it changes what the writer means
+ 최종목표	Your final result must be shorter than the original text in length and the meaning consistent with the initial meaning. Do your best!
최종목표	Your final result must be shorter than the original text in length and the meaning consistent with the initial meaning. Do your best!
표적형목표	Check out whether you removed... a) Overlapped words      b) Unnecessary sentences

## 5.5 측정

본 실험은 결과물의 질을 a) 작업길이, b) 작업시간, c) 작업비용 등 3가지로 측정했다. 이는 클라우드소싱이라는 작업환경과 텍스트 단축이라는 작업내용을 고려하여 설정된 척도이다. 클라우드소싱의 특징이자 장점 중 하나는 빠른 시간 안에 적은 비용으로 원하는 작업 결과물을 얻을 수 있다는 것이다. 그러므로 작업의 목적에 부합하면서도 작업 효율이 높은 결과물이 클라우드소싱에서는 좋은 결과물이라고 할 수 있다. 작업길이는 결과물의 질적 측면 중 작업의 목적 부합성을 측정한다. 짧은 시간이 소요되더라도 충분히 길이가 줄어들지 않았다면 이는 작업의 본래 목적에 다른 결과물이기 때문에 좋은 결과물이라 할 수 없다. 또 목적에 충실한 결과물을 얻었다고 하더라도 지나치게 많은 시간과 비용을 들이게 된다면, 이는 클라우드소싱에서 좋은 결과물이라 하기 어렵다. 따라서 작업시간과 작업비용은 결과물의 질적 측면 중 클라우드소싱 환경에서 중요한 작업의 효율 측면을 측정할 수 있다. 이처럼 작업길이, 작업시간, 작업비용 등 3가지 척도를 통해 작업 결과물의 질을 목적 부합성 및 효율 측면에서 파악하였다.

먼저 작업길이는 원래 텍스트 대비 최종 결과물의 길이가 얼마나 줄었는지 비율(%)로 측정한다. 텍스트에 포함된 글자(character) 수를 기준으로 최종적으로 제출된 답안의 길이가 얼마나 되는지 확인하여 피드백과 목표가 결과물에 영향을 주었는지 알아보기 위해서이다. 만약 주어진 텍스트가 100자이고, 참여자가 긍정적인 피드백 메시지를 받고 최종 제출한 결과물이 원래의 80%인 80자라면 그 메시지는 20%를 줄이는 만큼의 효과를 가진다고 볼 수 있다. 실험 중 참여자에게는 텍스트가 2가지 종류 중 임의로 배정되는데, 제공받은 텍스트의 길이에서 줄어든 비율을 측정하여 텍스트의 길이에 따른 영향력을 상쇄시키고자 한다.

다음으로 작업시간은 길이 1%를 줄이는데 필요한 시간으로 측정했는데, 총 소요된 시간을 줄어든 길이의 비율로 나누어 비교가 용이하도록 했다. 만약 100초의 작업시간이 소요되었고 제출된 결과물이 원 텍스트의 90% 길이라면 10%을 줄이는데 100초가 소요된 것이다. 따라서 해당 작업자는 10초/%라는 작업효율을 갖게 된다. 그러므로 특정 커뮤니케이션 조건에서 작업시간이 감소할 경우, 1%를 줄이는데 들이는 시간이 줄어든 것으로 작업 효율이 증가했다고 판단할 수 있다.

아울러 작업비용 또한 길이 1%를 줄이는데 필요한 비용으로, 총 비용을 줄어든 길이의 비율로 나누었다. 실험에서는 제시문의

길이를 1% 줄일 때마다 \$0.15 혹은 \$0.3을 지급했는데, 2단계의 답안을 기준으로 지급했기 때문에 1단계와 2단계 답안의 작업비용을 비교하여 작업 효율이 증가했는지 비교할 수 있다. 또한 서로 다른 커뮤니케이션 메시지를 받았을 때의 효율을 비교할 수도 있다. 그러므로 작업비용은 이처럼 길이 1%를 줄이기 위한 비용으로 조작화하여 측정하였다.

## 6. 연구결과

### 6.1 데이터 클리닝

실험으로 수집된 데이터 중 결과 분석에 적합하지 않은 데이터를 선별해내기 위해 일정한 절차를 거쳐 데이터 클리닝을 실시했다. 본 실험의 내용이 제시문의 의도를 훼손하지 않는 범위 내에서 길이를 줄이는 작업이기 때문에 제출된 답안이 요구사항대로 작성되었는지 파악하는 작업이 필요했다. 따라서 연구자는 참여자들이 작성한 모든 답안들을 검토하였으며, 이 중 적절하지 못한 결과물을 찾아 분석에서 제외하였다. 적절성 여부는 2가지 기준으로 판단하였는데 사전에 연구자가 정의한 키워드를 얼마나 가지고 있는지를 통해 원 텍스트의 의도를 해치지 않았는지 파악했다. 또한 답안 내용과 더불어 작성 시간 및 길이를 참고해 참여자가 성실하게 작성했는지를 판단하였다. 이 두 가지 기준을

충족시키고 최종 분석 대상이 된 데이터는 모두 제시문의 요약문으로서 적절한 수준을 충족했다고 볼 수 있다. 그러므로 그 다음 단계에서 종속변인들을 실험 조건별로 비교하는데 큰 무리가 없다고 판단하였다.

표 6. 키워드 리스트

텍스트 번호	키워드
1	Graphical user interface(GUI), the GUI approach, the rich interface modalities, the GUI paradigm, world-situated devices, Tangible User Interfaces(TUIs), the metaDESK system, TangibleGeospace, Tangible Bits Project, the transBOARD, ambientROOM, physical objects, ambient environmental displays
2	Print publishers, Apple's new iPad, digital editions, interactive graphics, magazine cover, the background image, video loop, new experiences, readers

먼저 참여자가 제출한 답안이 키워드 중 절반 이상 포함하지 못한 경우에는 부적절한 답안으로 판단했다. 원 텍스트 중 1번째 텍스트는 HCI 분야의 논문 중에서 발췌한 것으로, 기존 GUI 방식이 사용자와 물리적 환경을 연결하는데 한계가 있어 새로운 방식인 metaDESK 를 제안한다는 것이 중요 내용이다. 따라서 기존 방식에 대한 설명이나 metaDESK 에 대한 설명이 지나치게 삭제된 경우 부적절한 결과물로 판단하였다. 원 텍스트 중 2번째 텍스트는

애플(Apple)사의 아이패드(iPad) 등장 후 잡지 표지가 사람들의 이목을 끌기 위해 다양한 인터페이스를 제공하는 쪽으로 바뀌고 있다는 내용이다. 따라서 아이패드 등장과 잡지 표지의 관계에 대한 내용이나 새로운 잡지 표지에 대한 설명이 지나치게 삭제된 경우 부적절하다고 판단했다.

아울러 작업자의 관여도가 현저히 낮다고 판단되는 경우도 적절하지 않은 결과물로 분류했다. 내용과 작업시간 및 길이를 모두 고려하여 판단했는데, 내용 측면에서 관여도가 낮은 경우의 답안의 예로는 1번째 단계에서는 주어진 텍스트의 반절을 적고 2번째 단계에서는 나머지 반절을 적는 경우 등이 있다. 또 단순히 길이를 줄여 제출하기 위해 한두 단어를 삭제한 경우 등의 경우에도 제외했다. 아울러 제한시간 10분 중 작업시간이 지나치게 짧고 길이도 10% 이내로 수정된 경우 등 참여자가 충분히 고려하지 않은 답안으로 판단된 경우에도 답안을 제외하였다. 이처럼 2가지 기준을 통해 적절하지 않은 답안을 제외하여 총 344명의 참여자들의 제출한 답안들을 최종분석의 결과물로 선정하였다.

## 6.2 기술통계

데이터 클리닝 절차에 따라 최종분석에는 총 344명의 데이터가 사용되었다. 구체적으로 살펴보면 보수 조건에 따라 \$0.15 조건은 165명, \$0.3 조건은 179명이었다. 보수 조건별로 각각 8개의 조건에서 참여자를 모집하였고, 참여자 등록 시 랜덤하게 조건을 부여하는 기능으로 조건별 표본 수를 동일 혹은 유사하게 하였다. 조건별로 참여자 수는 <표 7>과 같으며 종속변인의 기술통계는 <표 8> 및 <표 9>와 같다.

표 7. 조건별 참여자 수

보수	긍정적 피드백	부정적 피드백	서술적 피드백	평가적 피드백	단계목표+ 최종목표	최종목표	표적형 목표	성취형 목표	계
\$0.15	21	18	22	19	19	22	22	22	165
\$0.30	25	22	26	21	23	22	20	20	179
계	46	40	48	40	42	44	42	42	344

표 8. \$0.15 조건의 기술통계

조건 *	작업길이		작업시간		작업비용	
	Mean	Sd	Mean	Sd	Mean	Sd
긍정적피드백(1)	66.33	21.23	27.09	34.77	0.009	0.011
긍정적피드백(2)	63.57	19.27	14.34	14.34	0.006	0.005
부정적피드백(1)	71.67	17.69	23.31	19.38	0.010	0.010



부정적피드백(2)	62.22	23.29	11.78	11.58	0.006	0.004
서술적피드백(1)	59.86	18.43	16.89	13.48	0.005	0.003
서술적피드백(2)	50.32	18.91	10.49	9.35	0.004	0.002
평가적피드백(1)	58.05	18.37	14.90	11.96	0.005	0.003
평가적피드백(2)	39.47	17.42	6.42	6.52	0.003	0.001
단계+최종목표(1)	59.47	23.30	18.47	16.97	0.006	0.005
단계+최종목표(2)	53.37	23.27	10.35	12.57	0.004	0.003
최종목표(1)	67.14	18.88	21.86	17.06	0.007	0.004
최종목표(2)	59.50	18.19	12.23	12.43	0.005	0.003
표적형목표(1)	58.31	23.86	17.61	21.27	0.007	0.008
표적형목표(2)	50.91	23.14	8.92	12.07	0.004	0.004
성취형목표(1)	59.59	22.39	16.66	12.58	0.005	0.004
성취형목표(2)	53.09	21.86	10.57	9.80	0.004	0.003
* 조건(1): 해당 조건의 1단계 / 조건(2): 해당 조건의 2단계						

표 9. \$0.3 조건의 기술통계

조건 *	작업길이		작업시간		작업비용	
	Mean	Sd	Mean	Sd	Mean	Sd
긍정적피드백(1)	71.04	18.30	29.69	39.49	0.020	0.023

긍정적피드백(2)	66.30	17.84	18.89	37.66	0.016	0.021
부정적피드백(1)	67.62	27.36	31.24	33.49	0.030	0.063
부정적피드백(2)	60.43	22.69	18.44	33.88	0.024	0.064
서술적피드백(1)	62.31	25.67	21.06	32.71	8.48	5.09
서술적피드백(2)	58.35	24.16	10.51	19.22	9.42	4.63
평가적피드백(1)	66.41	23.25	29.06	29.16	0.019	0.022
평가적피드백(2)	52.09	24.12	14.44	15.25	0.009	0.008
단계+최종목표(1)	67.43	23.84	27.79	27.29	0.016	0.013
단계+최종목표(2)	59.00	24.27	13.63	14.55	0.011	0.007
최종목표(1)	56.78	13.60	7.92	8.12	0.008	0.002
최종목표(2)	41.61	12.63	3.81	4.31	0.005	0.002
표적형목표(1)	60.70	19.80	17.68	11.56	0.010	0.006
표적형목표(2)	54.60	18.78	8.88	5.88	0.008	0.005
성취형목표(1)	65.85	15.92	18.12	11.35	0.011	0.006
성취형목표(2)	54.80	19.70	10.58	6.95	0.008	0.004

\* 조건(1): 해당 조건의 1단계 / 조건(2): 해당 조건의 2단계

## 6.3 세부 연구결과

여기서부터는 본 연구가 앞서 제시했던 가설을 검증하고, 크라우드소싱의 결과물이 어떤 조건에서 향상되는지 살펴보고자 한다. 실험조건별 종속변인을 비교하기 전, 데이터 전체를 대상으로 커뮤니케이션과 보수의 양이 결과물의 질에 미치는 영향력을 종합적으로 파악하고자 이원 분산분석과 회귀분석을 실시하였다. 그 다음 커뮤니케이션 조건별로 1단계와 2단계 답안간 대응표본 t 검정을 실시하고, 참여자에게 제공된 커뮤니케이션이 작업자의 행동 변화를 야기했는지 알아보았다. 이어 조건 간에 t 검정을 이용하여 종속변인의 차이를 파악했다. 커뮤니케이션 조건들이 종속변인을 얼마나 증가 혹은 감소시켰는지 그 차이의 크기를 조건별로 비교했다. 마지막으로 보수 조건과 커뮤니케이션 조건을 모두 조합하여 동일한 커뮤니케이션 조건 및 단계에서 보수 조건이 결과물의 질에 영향을 미쳤는지 파악했다.

### 6.3.1 커뮤니케이션과 보수의 양의 영향력 비교

본 연구는 실험 참여자에게 8개의 커뮤니케이션 및 2개의 보수 조건을 부여하고 각 조건에서 작업길이, 작업시간 및 작업비용을 측정하였다. 본 연구의 가장 큰 목적은 작업자와의 커뮤니케이션이 작업 결과물의 질적 향상에 미치는 영향에 대해 탐구하는 것이지만,

균형있는 시각을 위해 기존 논의에서 비중있게 다뤄진 보수적 양의 갖는 영향력에 대해 함께 살펴볼 필요가 있다. 따라서 가설 검증에 위한 세부적인 분석에 앞서, 커뮤니케이션과 보수조건에 따른 결과물의 질의 차이가 있는지 파악하였다. 이때 전체 결과물의 질을 판단하기 위해 작업길이, 작업시간 및 작업비용을 합산한 새로운 변인인 통합 품질 변인을 추가하였다. 본 연구의 종속변인인 작업길이와 작업시간은 짧을수록, 그리고 작업비용은 적을수록 결과물의 질이 높은 것으로 볼 수 있다. 이처럼 세 종속변인 모두 값이 작아질 수록 긍정적인 결과물로 판단되므로, 이들을 합산하여 통합 품질 변인을 만든다면 그 수치가 낮을수록 긍정적이라고 볼 수 있다. 이에 따라 커뮤니케이션 조건과 보수 조건을 두 요인으로 하고, 종속변수를 통합 품질로 하여 이원분산분석을 실시하였다.

〈표 10〉은 각 조건에 해당되는 통합 품질의 평균값을 정리한 것이고, 〈표 11〉은 커뮤니케이션 및 보수 조건에 따른 통합 품질을 살펴본 결과이다. 이원 분산분석 결과에 따르면 통합 품질에 커뮤니케이션 조건만이 유의미한 주효과를 가지고 있었으며( $p < .001$ ), 보수 조건은 그렇지 못했다. 또한 커뮤니케이션 및 보수 조건의 상호작용 효과도 유의미한 수준이 아니었다. 이처럼 커뮤니케이션에 따라 통합 품질의 차이가 있을 뿐, 보수 조건에

따른 차이는 발견하지 못했다. 그러므로 실험 결과 전반에 대해 커뮤니케이션 조건이 결과물의 질과 연관되어 있음을 알 수 있었다.

표 10. 통합 품질 기술통계

보수 조건	커뮤니케이션 조건								
		1	2	3	4	5	6	7	8
	1	85.68	81.13	68.79	59.42	69.04	79.32	67.89	74.38
2	92.42	85.32	75.45	80.13	82.47	60.62	70.94	72.88	

\* 통합 품질 = 작업길이 + 작업시간 + 작업비용

\*\* 통합 품질 수치가 낮을 수록 긍정적인 결과임

표 11. 이원분산분석 결과표\_ 커뮤니케이션 x 보수

요인	자유도	제곱합	평균제곱	F	유의확률
보수	1	3302	3302	2.242	0.134697
커뮤니케이션	1	17719	17719	12.035	0.000553
보수 x 커뮤니케이션	1	3398	3398	2.308	0.129162
잔차	734	1080683	1472		

이원 분산분석에 이어 커뮤니케이션 및 보수 조건을 터미 변수로 하고, 통합 품질 변인을 종속변인으로 하여 회귀분석을 실시하였다. 그 결과는 아래 <표 12>와 같다. 커뮤니케이션 조건은 부정적 피드백을 제외하고 모두 통합 품질에 유의미한 영향력을 가지고 있었으나, 보수 조건은 그렇지 못했다. 이처럼 보수의 양보다는 커뮤니케이션 방식이 결과물의 질에 보다 큰 영향력을 가지는

것으로 나타나, 보수의 양적 증가가 결과물의 질적 향상에 절대적인 영향력을 발휘하지는 않는다는 점을 알 수 있었다.

또, 통합 품질 변인은 그 수치가 낮을수록 결과물의 질이 높은 것을 의미하므로 독립변인의 회귀계수가 음의 값일 때 해당 변인이 결과물의 질을 향상시킨다고 볼 수 있다. 회귀분석 결과 커뮤니케이션 조건의 회귀계수는 모두 음의 값을 가져, 커뮤니케이션은 통합 품질 변인과 유의미한 관계일 뿐만 아니라 실질적으로 결과물의 질을 높이는 방향으로 영향력을 행사하고 있었다. 반면 보수 조건은 그렇지 못했다. 그러므로 결과물의 질에는 보수 조건보다 커뮤니케이션이 긍정적인 영향력을 준다고 볼 수 있었다.

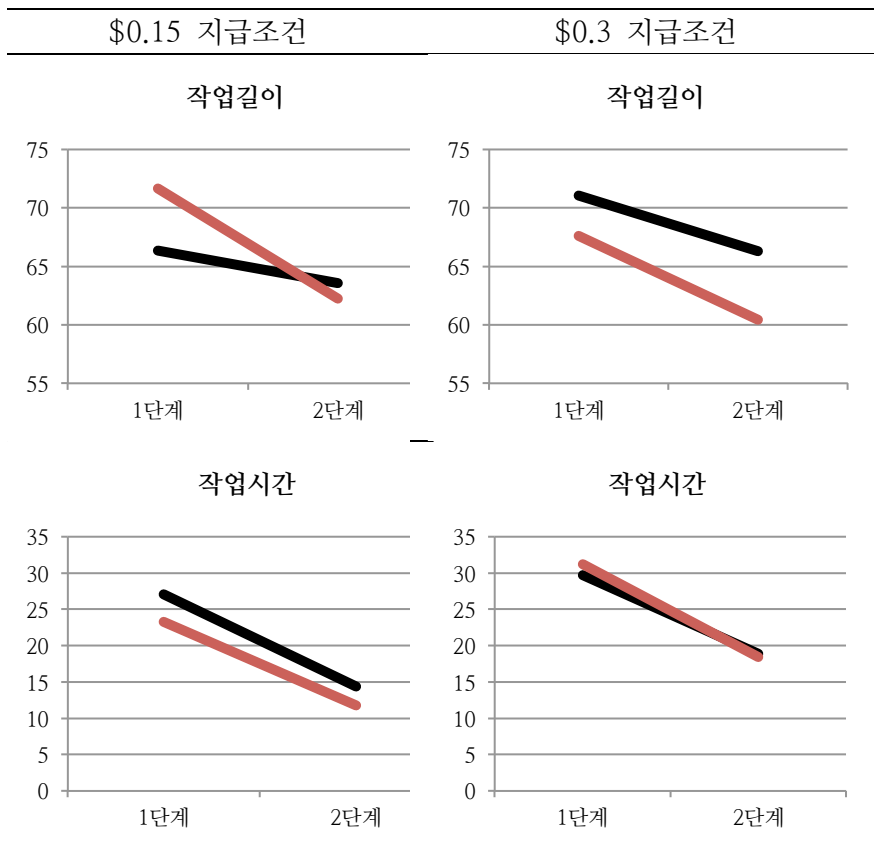
이원 분산분석 및 회귀분석 결과 보수의 양보다는 커뮤니케이션이 결과물의 질적 향상에 보다 효과적이라는 점을 알 수 있었다. 그러나 이는 실험 데이터 전반에 대한 분석이므로, 보다 세밀하게 살펴볼 필요가 있다. 그러므로 다음 단락부터는 가설 검증을 위해 비교 대상이 되는 커뮤니케이션 조건끼리 결과물의 질에 미치는 영향력을 비교할 것이다.

표 12. 회귀분석 결과

	회귀계수	표준오차	t	p
상수	87.17	4.25	20.52	$p < .001$
\$0.30 지급	4.04	2.83	1.43	0.153

부정적피드백	-5.97	5.68	-1.05	0.294
서술적피드백	-16.92	5.50	-3.08	$p < .05$
평가적피드백	-18.53	5.73	-3.23	$p < .05$
단계목표+최종목표	-13.29	5.54	-2.40	$p < .01$
최종목표	-19.80	5.52	-3.59	$p < .001$
표적형목표	-19.75	5.75	-3.43	$p < .001$
성취형목표	-15.47	5.64	-2.75	$p < .01$

### 6.3.2 긍정적 피드백과 부정적 피드백의 효과 비교



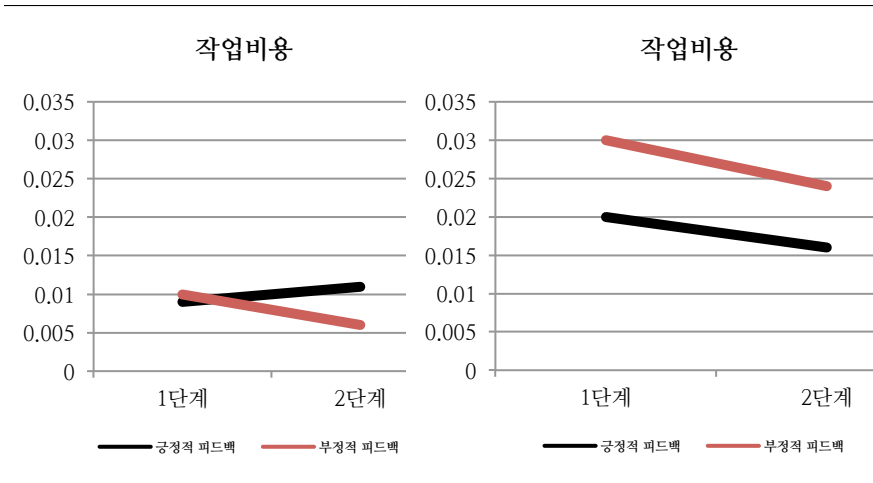


그림 10. 긍정적 피드백과 부정적 피드백의 단계별 종속변인

첫 번째 가설과 관련하여 먼저 긍정적 피드백과 부정적 피드백이 실험 과정에서 참여자의 행동 변화를 일으켰는지 살펴보았다. 긍정적 피드백을 제공하는 경우, \$0.15를 지급하는 조건에서 긍정적 피드백은 작업길이와 시간을 감소시켰고 비용에는 변화가 없었다. 작업길이는 1단계( $M=66.33$ ,  $SD=21.23$ )보다 2단계( $M=63.57$ ,  $SD=19.27$ )에서 유의미한 차이를 보였다( $t(20)=2.52$ ,  $p < .05$ ). 작업시간도 1단계( $M=27.10$ ,  $SD=34.77$ )와 2단계( $M=14.34$ ,  $SD=14.34$ )로 유의한 차이를 나타냈으나( $t(20)=2.36$ ,  $p < .05$ ), 작업비용은 차이가 나타나지 않았다. 즉 긍정적 피드백을 받은 후에는 작업 목적 부합성 및 시간 효율성이 향상되었고, 비용 효율성에는 유의미한 변화가 없었다. 같은 분석을 \$0.3 지급 조건에서도 실시한 결과 세 가지 종속변인 모두에서 유의한 변화가 발견되었다. 긍정적 피드백 조건에서 1단계



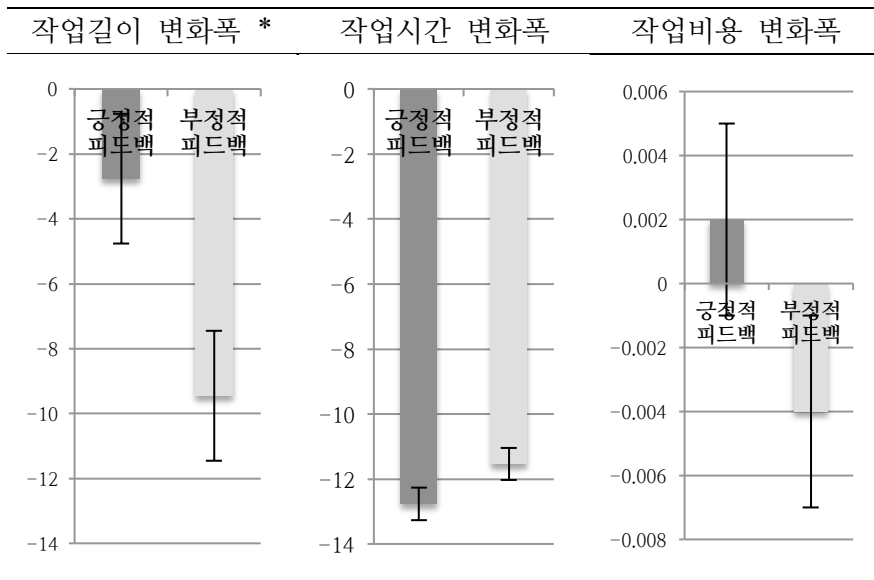
작업길이( $M=71.04$ ,  $SD=18.30$ )는 2단계( $M=68.30$ ,  $SD=17.84$ )에서 유의미하게 감소했고( $t(22)= 5.49$ ,  $p < .05$ ), 작업시간도 1단계( $M=29.69$ ,  $SD=39.49$ )에서 2단계( $M=18.89$ ,  $SD=37.66$ )로 진행하며 유의하게 줄어들었다( $t(22)=7.22$ ,  $p < .05$ ). 또한 1단계 답안에서 요구된 비용( $M=0.020$ ,  $SD=0.023$ )은 2단계( $M=0.016$ ,  $SD=0.021$ )에서 유의하게 감소했다( $t(22)= 3.18$ ,  $p < .05$ ). 이처럼 \$0.15 지급 시와 달리 \$0.3 지급 시에는 긍정적 피드백 메시지를 받은 후 작업 목적 부합성과 시간 및 비용 효율성이 향상되었다. 이처럼 단계간 종속변인을 비교했을 때 긍정적 피드백은 작업 결과물의 질을 높이고 있었다.

이어 부정적 피드백을 제시한 경우 \$0.15 지급 조건에서 작업길이는 1단계( $M=71.67$ ,  $SD=17.69$ )와 2단계( $M=62.22$ ,  $SD=23.29$ ) 간의 유의한 차이가 나타났고( $t(17)=4.08$ ,  $p < .05$ ), 마찬가지로 작업시간도 1단계( $M=23.31$ ,  $SD=11.78$ )보다 2단계( $M=11.78$ ,  $SD=11.58$ )에서 감소했다( $t(17)= 4.44$ ,  $p < .05$ ). 아울러 긍정적 피드백 조건과 다르게 부정적 피드백 조건의 1단계( $M=0.01$ ,  $SD=0.010$ )의 비용은 2단계( $M=0.006$ ,  $SD=0.004$ )에서 유의하게 감소했다( $t(17)= 2.20$ ,  $p < .05$ ). 즉 \$0.15를 지급하는 조건에서 부정적 피드백은 작업단계가 진행됨에 따라 작업의 목적 부합성과 시간 및 비용 효율성을 모두 향상시켰으며, 이는 비용 효율성에 유의한 차이를 보이지 않은

긍정적 피드백 조건과 대비된다. 보수조건 \$0.3에서 동일한 검정을 실시한 결과, 부정적 피드백을 제시했을 때 작업길이는 1단계( $M=67.62$ ,  $SD=27.36$ )보다 2단계( $M=60.43$ ,  $SD=22.69$ )에서 유의하게 줄었고( $t(20)=2.75$ ,  $p < .05$ ), 작업시간도 1단계( $M=31.24$ ,  $SD=33.49$ )와 2단계( $M=18.43$ ,  $SD=33.88$ ) 간에 차이가 나타났다( $t(20)=3.36$ ,  $p < .05$ ). 작업비용도 1단계( $M=0.030$ ,  $SD=0.063$ )에서 2단계( $M=0.024$ ,  $SD=0.064$ )로 진행하면서 감소하는 양상을 보였다( $t(20)=3.76$ ,  $p < .05$ ). 이러한 결과를 종합해보면 \$0.15 지급 조건에서와 마찬가지로 부정적 피드백 메시지가 작업자에게 제공된 후에는 작업길이, 작업시간, 작업비용이 모두 감소하였다. 따라서 목적에 부합하면서도 시간과 비용을 효율적으로 활용한 결과물을 도출할 수 있었다.

위의 분석에서 긍정적 피드백과 부정적 피드백은 참여자들의 행동을 변화시키고 작업 결과물의 질을 향상시키는 영향력을 가지고 있음을 알 수 있었다. 그 다음 단계로 두 피드백이 결과물의 질적 향상에 미치는 영향력을 비교하기 위해 종속변인의 변화폭에 대한  $t$  검정을 실시했다. 작업길이를 긍정적 피드백과 부정적 피드백이 변화시킨 정도를 비교한 결과, \$0.15를 지급하는 조건에서는 부정적 피드백이 긍정적 피드백보다 작업길이를 감소시키는 영향력이 발견되었다. 긍정적 피드백을 받은 참여자들이 텍스트를 줄인 폭( $M=-2.76$ ,  $SD=5.02$ )과 부정적

피드백을 받은 참여자들이 줄인 폭( $M=-9.44$ ,  $SD=9.82$ ) 간에는 유의한 차이가 있었다( $t= 2.61$ ,  $df= 24.43$ ,  $p < .05$ ). 그러나 \$0.3을 지급하는 경우에는 두 피드백의 영향력에 유의한 차이가 발견되지 않았다. 그러나 변화의 방향은 \$0.15 지급 시와 동일하게 나타나 부정적 피드백이 긍정적 피드백에 비해 결과물의 질을 향상시키는 것으로 나타났다. 긍정적 피드백에서 작업길이의 변화폭( $M=-4.74$ ,  $SD=4.14$ )보다 부정적 피드백에서의 변화폭( $M=-7.19$ ,  $SD=11.97$ )이 더 컸기 때문이다. 또한 작업시간과 비용의 변화폭도 \$0.15 지급 조건과 마찬가지로 유의미한 차이는 없었지만 변화 방향이 같았으므로 동일한 효과를 나타낸다고 해석할 수 있다. 아울러 동일한 커뮤니케이션 조건 및 단계에서 보수의 양이 다를 때, 종속변인의 변화 폭이 달라지는지 살펴보았다. 그 결과 모두 유의미한 차이가 없어, 긍정적 혹은 부정적 피드백을 제공했을 때 보수의 양은 결과물의 질적 향상에 별다른 영향력을 미치지 않는 것으로 나타났다.



\*  $p < .05$

그림 11. 긍정적 피드백 vs. 부정적 피드백\_종속변인 변화폭(\$0.15)

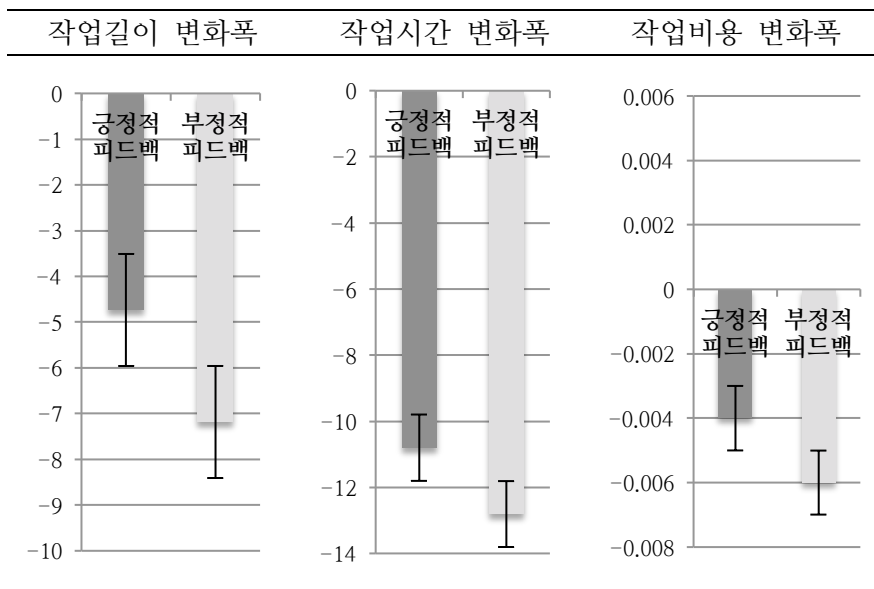


그림 12. 긍정적 피드백 vs. 부정적 피드백\_종속변인 변화폭(\$0.3)

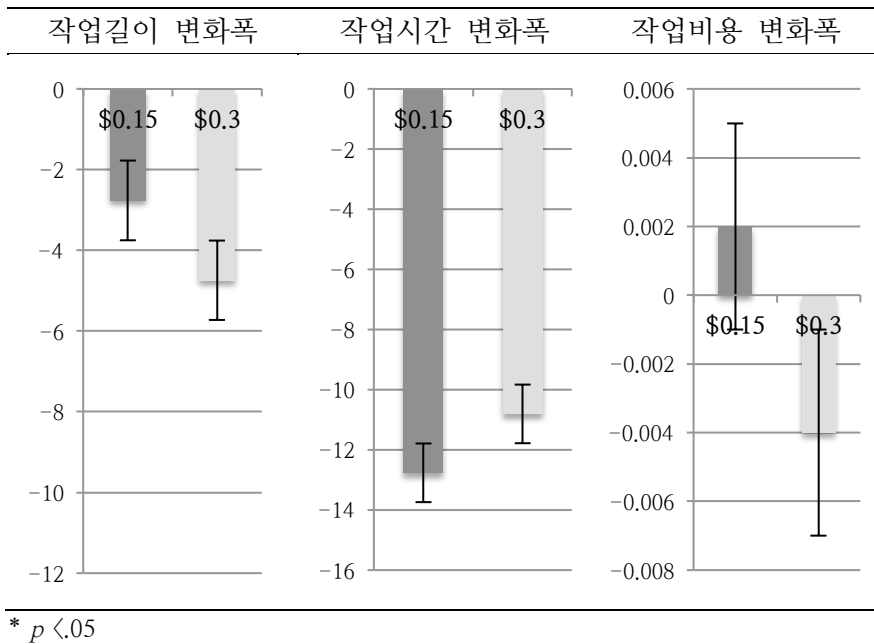


그림 13. 긍정적 피드백\_보수조건별 종속변인 변화폭

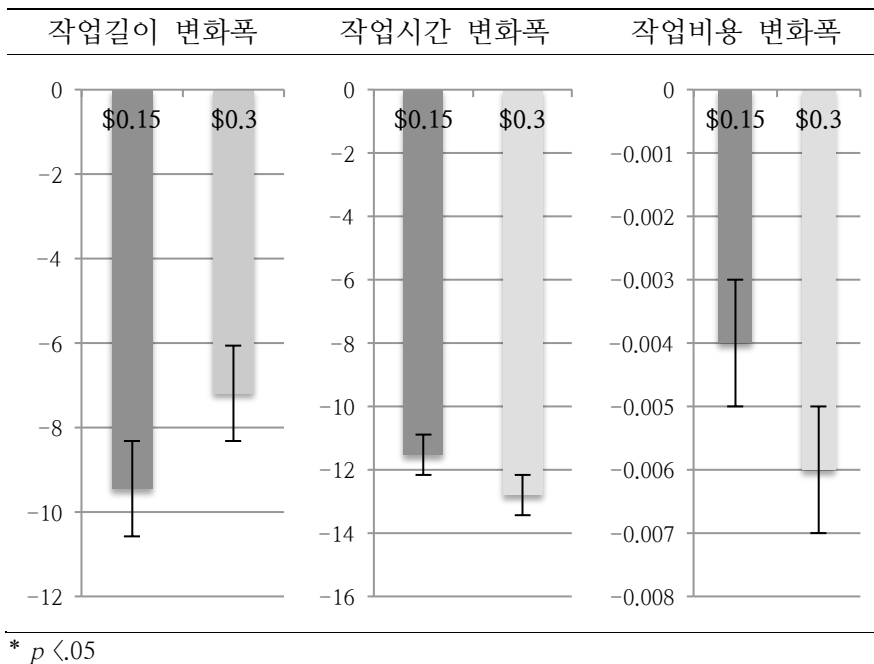


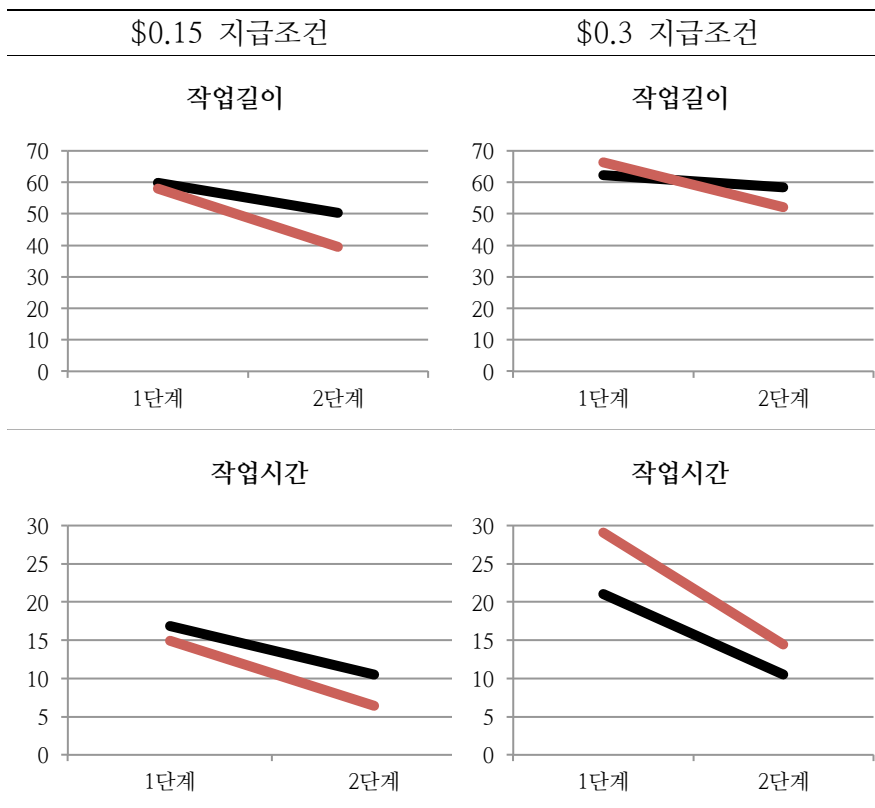
그림 14. 부정적 피드백\_보수조건별 종속변인 변화폭

위와 같은 결과를 종합하면 긍정적 피드백과 부정적 피드백은 모두 1단계에서 2단계로 진행하면서 작업길이, 시간 및 비용을 감소시켜 참여자의 행동 변화를 유도함을 알 수 있었다. 그리고 두 피드백이 결과물의 질적 향상에 미치는 영향력을 비교한 결과 \$0.15를 지급하는 조건에서 부정적 피드백이 긍정적 피드백과 비교하여 작업길이를 유의미하게 줄였고, \$0.3을 지급할 때는 그러한 차이가 통계적으로 유의미하지는 않았지만 영향력의 방향은 동일하게 나타났다. 이는 긍정적 피드백보다 부정적 피드백이 결과물의 질적 향상에 효과적이라는 가설 1.1을 뒷받침한다. 부정적 피드백은 긍정적 피드백과 작업시간과 비용에서 차이가 없는 상태에서 작업길이를 감소시켰으므로, 보다 효율적으로 작업의 목적에 부합하는 결과물을 도출했기 때문이다. 그러므로 클라우드소싱에서 부정적 피드백은 긍정적 피드백보다 작업의 목적에 부합하는 결과물을 효율적으로 도출한다고 볼 수 있다.

두 피드백의 영향력에 더해 보수의 양이 갖는 영향력도 생각해볼 필요가 있다. 보수의 양은 결과물의 질을 높이기 위해 사용되는 대표적인 방법이므로 그 영향력을 커뮤니케이션 조건의 영향력과 함께 파악해 보았다. 그 결과 절대적인 수치는 보수의 양이 많을 때 높은 경우도 있었으나 유의미한 차이는 나타나지 않았다. 반면 \$0.15 지급 시 긍정적 피드백보다 부정적 피드백이 작업길이를 유의미하게 줄여 커뮤니케이션 조건 간 영향력의 차이는 보다

뚜렷하게 나타나고 있었다. 이러한 결과를 통해 보수의 양이 결과물의 질을 향상시키는데 절대적인 영향력을 가지고 있지는 않으며, 커뮤니케이션도 결과물의 질을 높이는데 기여할 수 있다는 본 연구의 주장을 뒷받침할 수 있었다.

### 6.3.3 서술적 피드백과 평가적 피드백의 효과 비교



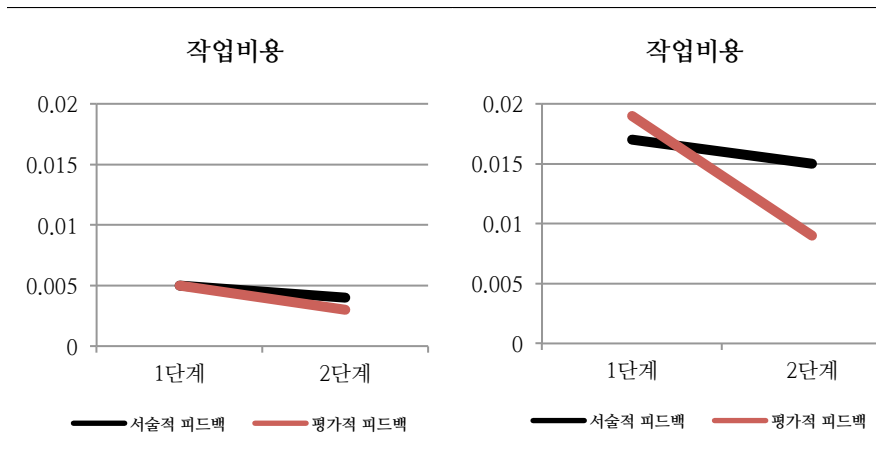


그림 15. 서술적 피드백과 평가적 피드백의 단계별 종속변인

서술적 피드백과 평가적 피드백이 결과물의 질에 어떤 영향력을 가지고 있는지 살펴보기 위해 대응표본 t 검정을 실시하였다. 그 결과 두 가지 피드백 모두 \$0.15를 지급하는 경우와 \$0.3을 지급하는 경우에서 작업길이, 시간, 비용 등 3가지 종속변인에 유의미한 차이를 가져왔다. 먼저 서술적 피드백을 제공했을 때 \$0.15 지급 조건에서 작업길이는 1단계( $M=59.86$ ,  $SD=18.43$ )보다 2단계( $M=50.32$ ,  $SD=18.91$ )에서 유의미하게 감소했고( $t(21)=5.54$ ,  $p < .05$ ), 작업시간도 1단계( $M=16.89$ ,  $SD=13.48$ )와 2단계( $M=10.50$ ,  $SD=9.35$ ) 간에 유의한 차이가 나타났다( $t(21)=4.37$ ,  $p < .05$ ). 작업비용도 1단계( $M=0.005$ ,  $SD=0.003$ )보다 2단계( $M=0.004$ ,  $SD=0.002$ )가 유의하게 적었다( $t(21)=3.56$ ,  $p < .05$ ). 즉, \$0.15 지급 시에 서술적 피드백 메시지를 받은 이후의 작업길이와 시간 및 비용은 모두 메시지를 받기 전보다 감소하여 결과물이 질적으로



향상되었다. 이러한 추이는 \$0.3 지급 조건에서도 동일하게 나타났는데, 작업길이는 1단계( $M=62.31$ ,  $SD=25.67$ )보다 2단계( $M=58.35$ ,  $SD=24.16$ )에서 유의하게 감소했고( $t(25)=3.37$ ,  $p < .05$ ), 작업시간도 1단계( $M=21.06$ ,  $SD=32.71$ )보다 2단계( $M=10.51$ ,  $SD=19.22$ )에서 유의한 차이가 나타났다( $t(25)=3.67$ ,  $p < .05$ ). 아울러 작업비용도 1단계( $M=0.017$ ,  $SD=0.028$ )에서 2단계( $M=0.015$ ,  $SD=0.028$ )로 넘어가면서 감소했다( $t(25)=3.36$ ,  $p < .05$ ). 이와 같이 서술적 피드백은 두 가지 보수조건에서 모두 작업 목적 부합성, 시간 효율성, 그리고 비용 효율성을 향상시키는 영향력을 보여주었다.

평가적 피드백에 대해서도 대응표본 t 검정을 실시했는데, 서술적 피드백과 마찬가지로 모든 조건에서 종속변인에 유의미한 차이가 있었다. \$0.15를 지급하는 조건에서 작업길이는 1단계( $M=58.05$ ,  $SD=18.37$ )보다 2단계( $M=39.47$ ,  $SD=17.43$ )에서 유의하게 감소했고( $t(18)=6.31$ ,  $p < .05$ ), 작업시간은 1단계( $M=14.90$ ,  $SD=11.96$ )보다 2단계( $M=6.42$ ,  $SD=6.52$ )에서 유의한 차이를 보였다( $t(18)=5.55$ ,  $p < .05$ ). 작업비용도 1단계( $M=0.005$ ,  $SD=0.003$ )보다 2단계( $M=0.003$ ,  $SD=0.001$ )에서 유의하게 감소했다( $t(18)=4.49$ ,  $p < .05$ ). 서술적 피드백과 같이 해당 보수조건에서 평가적 피드백도 작업길이, 작업시간, 작업비용을 1단계에 비해 2단계에서 줄이는 효과를 가지고 있었다. 이는

\$0.3을 지급하는 조건에서도 동일하게 관찰되었다. 작업길이가 1단계( $M=66.41$ ,  $SD=23.25$ )보다 2단계( $M=52.10$ ,  $SD=24.12$ )에서 감소했고( $t(21)=7.56$ ,  $p < .05$ ), 작업시간이 1단계( $M=29.06$ ,  $SD=29.16$ )보다 2단계( $M=14.44$ ,  $SD=15.25$ )에서 줄었으며( $t(21)=4.50$ ,  $p < .05$ ), 작업비용 또한 1단계( $M=0.019$ ,  $SD=0.022$ )보다 2단계( $M=0.009$ ,  $SD=0.008$ )에서 감소했다( $t(21)=2.99$ ,  $p < .05$ ). 이와 같은 비교 결과를 종합해보면, 평가적 피드백과 서술적 피드백 조건 모두에서 작업 목적에 부합하고 시간 및 비용 효율성이 높은 결과물이 도출되었으며 보수 조건에 따른 차이는 없었다. 따라서 두 피드백은 작업 결과물의 질을 향상시키는 영향력을 가지고 있으며, 보수의 양에 큰 영향력을 받지 않는다고 볼 수 있었다.

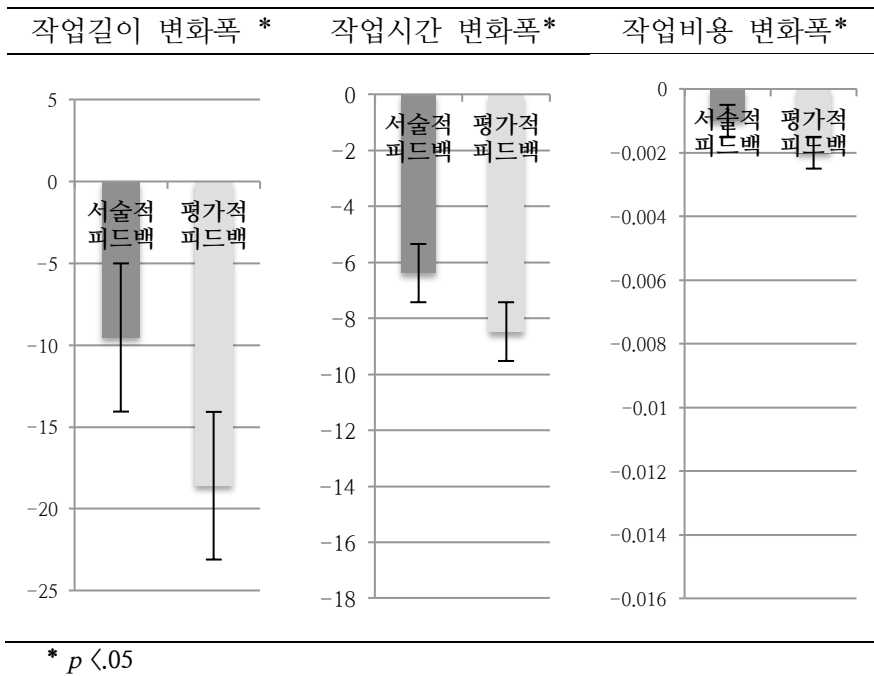


그림 16. 서술적 피드백 vs. 평가적 피드백\_종속변인 변화폭(\$0.15)

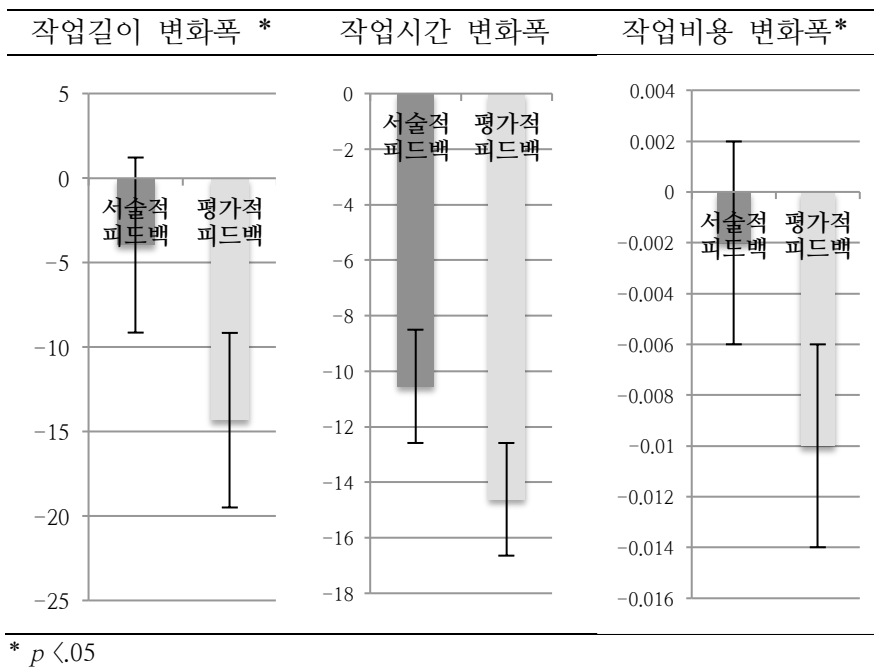


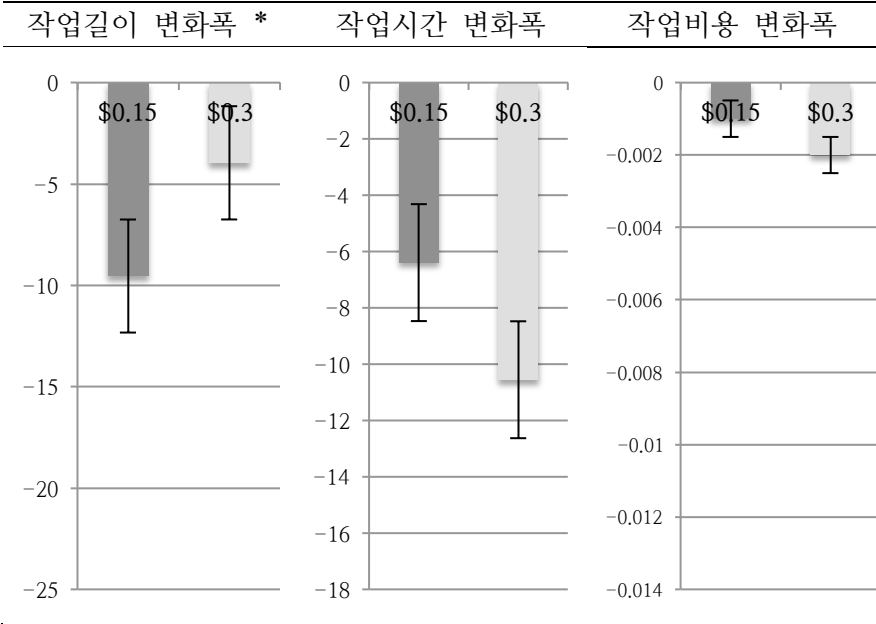
그림 17. 서술적 피드백 vs. 평가적 피드백\_종속변인 변화폭(\$0.3)

다음으로 서술적 피드백과 평가적 피드백이 결과물의 질에 미치는 영향력을 비교하기 위해 각 피드백이 작업길이, 시간, 비용을 변화시킨 폭의 차이를 살펴보았다. \$0.15 지급 조건에서 t 검정을 실시한 결과, 평가적 피드백을 제공한 경우의 작업길이 감소폭( $M=-18.58$ ,  $SD=8.08$ )은 서술적 피드백을 제공했을 때의 감소폭( $M=-9.55$ ,  $SD=12.83$ )과 유의한 차이를 보였다( $t(29.462)=2.65$ ,  $p < .05$ ). 반면 작업시간과 비용에는 유의미한 차이가 발견되지 않았다. 따라서 \$0.15 지급 시 평가적 피드백은 서술적 피드백보다 작업 목적 부합성, 시간 효율성, 비용 효율성 측면에서 더 좋은 결과물을 도출한다고 볼 수 있었다. \$0.3을 지급하는 조건에서는 평가적 피드백이 작업길이와 비용을 줄이는 효과가 발견되었다. 평가적 피드백을 제시했을 때 작업길이 감소폭( $M=-14.32$ ,  $SD=6.00$ )은 서술적 피드백을 제시한 경우의 감소폭( $M=-3.96$ ,  $SD=8.88$ )보다 유의하게 컸다( $t(35.84)=4.65$ ,  $p < .05$ ). 또한 작업비용은 평가적 피드백( $M=-0.009$ ,  $SD=0.014$ )이 서술적 피드백( $M=-0.002$ ,  $SD=0.003$ ) 조건에서보다 감소하는 폭이 유의미하게 컸다( $t(22.301)=2.36$ ,  $p < .05$ ). 반면 작업시간은 \$0.15 조건과 동일하게 서술적 피드백과 평가적 피드백에서 유의한 차이가 발견되지 않았다. 정리하면 \$0.3 지급 시에 평가적 피드백은 \$0.15 지급 조건처럼 작업시간을 줄이지는 못했지만, 작업길이와 작업비용을 서술적 피드백보다 줄였다. 그러나

절대값을 보면 평가적 피드백 조건에서 작업시간이 감소한 폭은 서술적 피드백 조건에서보다 커서, 효과의 방향은 동일하게 나타났다. 따라서 전반적으로 평가적 피드백이 서술적 피드백보다 작업 목적 부합성과 효율성 측면에서 우수한 결과물을 도출한다는 점을 발견할 수 있었다.

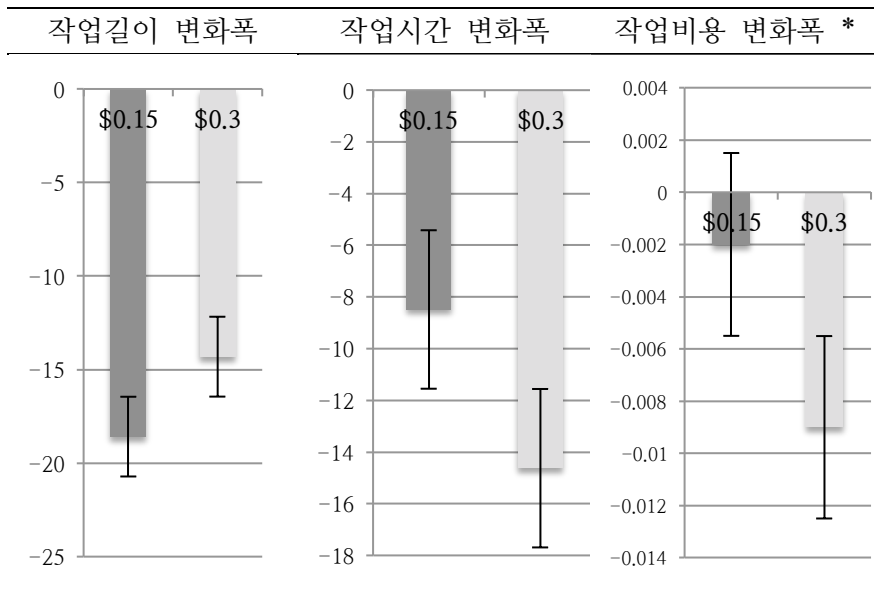
마지막으로 같은 커뮤니케이션 조건에서 보수의 양에 따라 종속변인의 변화 폭에 차이가 있는지 살펴보았다. 그 결과 서술적 피드백은 \$0.15를 지급하는 조건에서의 작업길이 감소폭( $M=-9.55$ ,  $SD=8.08$ )이 \$0.3을 지급하는 조건에서의 감소폭( $M=-3.96$ ,  $SD=6.00$ )과 유의한 차이를 보였다( $t(38,19)=-2.68$ ,  $p < .05$ ). 반면 평가적 피드백을 제공한 경우, \$0.15를 지급한 경우( $M=-0.002$ ,  $SD=0.002$ )보다 \$0.3을 지급한 경우( $M=-0.009$ ,  $SD=0.014$ )에 작업비용의 감소폭이 컸다( $t(21,73)=2.37$ ,  $p < .05$ ). 즉, 서술적 피드백을 제시했을 때는 \$0.15 지급 조건에서 보다 작업 목적에 부합하는 결과물이 도출되었고, 평가적 피드백을 제시했을 때는 \$0.3 지급 조건에서 보다 비용 효율성이 높은 결과물이 도출되었다. 아울러 유의미한 차이가 나타나지 않은 종속변인도 \$0.15를 지급할 때보다 \$0.3을 지급할 때 더 많이 감소하는 경우가 발견되었다. 그러므로 보수의 양은 기존 논의에서 주장한 바와 같이 결과물의 질에 일정한 영향을 끼치고 있음을 알 수 있었다. 그러나 보수의 양이 결과물의 질적 향상에 영향을 미친다 하더라도 절대적이지는

않았고, 영향력의 크기가 커뮤니케이션 조건에 따라 다르게 나타났다. 따라서 결과물의 질을 관리하기 위해서는 보수의 양 뿐만 아니라 커뮤니케이션 방식도 충분히 고려해야 할 필요가 발견되었다.



\*  $p < .05$

그림 18. 서술적 피드백\_ 보수조건별 종속변인 변화폭



\*  $p < .05$

그림 19. 평가적 피드백\_ 보수조건별 종속변인 변화폭

위와 같이 서술적 피드백과 평가적 피드백의 영향력을 비교 검증한 결과, 가설 1.2와 달리 평가적 피드백이 서술적 피드백보다 결과물의 질적 향상에 효과적인 것으로 나타났다. \$0.15 지급 조건에서는 평가적 피드백이 서술적 피드백보다 작업길이를 더 큰 폭으로 감소시켰고, \$0.3 지급 조건에서는 작업길이와 함께 비용도 큰 폭으로 감소시켰기 때문이다. 이는 선행연구에서 서술적 피드백이 평가적 피드백보다 결과의 향상에 효과적이라고 주장한 것과 상반된다. 해당 선행연구들은 대면 접촉이 있는 실제 근무 환경에서 이루어졌으므로 결과물이 좋고 나쁜 것을 평가하는 방식이 구성원들의 업무 의지에 악영향을 줄 수 있었다. 하지만 비대면 접촉이고 일회적인 관계로 이루어진 클라우드소싱에서는

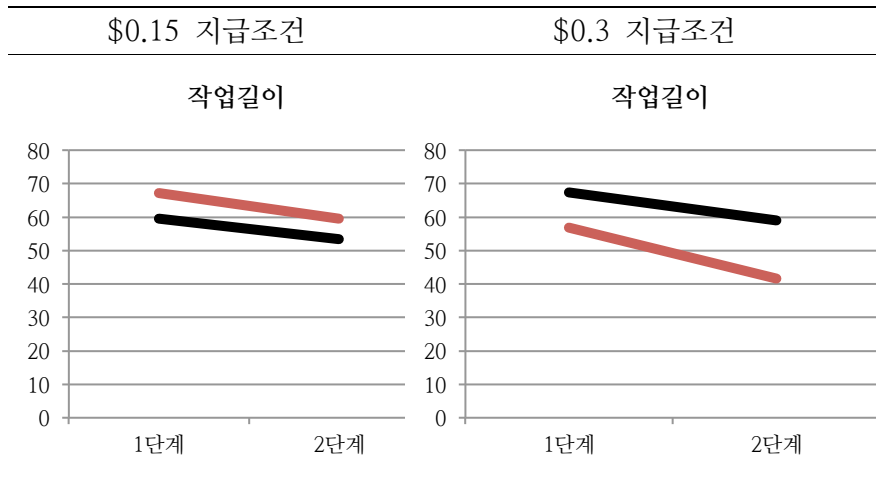
결과물에 대한 직접적인 판단이 명확한 판단근거를 작업자에게 제공했다고 해석할 수 있다. 그러므로 클라우드소싱에서는 평가적 피드백을 통해 작업자가 스스로의 작업을 즉시 판단할 수 있도록 돕는 것이 결과물의 질적 향상에 효과적임을 알 수 있었다.

또한 보수의 양이 미치는 영향력을 살펴보았는데, 긍정적 피드백 및 부정적 피드백을 비교했던 경우와 달리 그 영향력이 보다 뚜렷하게 나타났다. 절대적인 수치에서는 보수를 많이 주는 경우에 작업시간과 비용 등 효율과 관련된 종속변인이 더 많이 감소하는 것으로 나타났다. 그 중 평가적 피드백의 작업비용은 \$0.15를 지급할 때보다 \$0.3을 지급할 때 더 크게 감소하였다. 그러나 보수의 양을 늘리는 방식이 모든 측면에서 결과물의 질을 향상시키는 절대적인 방식이라고 보기는 어려웠다. 작업의 목적 부합성에는 오히려 반대의 영향력을 가지는 것으로 나타났기 때문이다. 서술적 피드백 조건에서는 \$0.15를 지급하는 경우 작업길이의 감소폭이 \$0.3을 지급하는 경우보다 크게 나타났고, 유의한 차이는 발견되지 않았더라도 동일한 방향의 영향력이 평가적 피드백 조건에서도 발견되었다. 또한 서술적 피드백과 평가적 피드백 조건에서 보수의 양이 미치는 영향력은 비슷한 방향으로 나타났지만, 그 크기가 다르게 나타나고 있었다. 즉, 보수의 양이 질적 향상에 미치는 영향력은 커뮤니케이션 조건에 따라 달랐다.



위와 같은 연구결과를 통해 볼 때, 보수를 늘리는 방식은 기존 논의와 마찬가지로 결과물의 질을 높였지만 작업자와의 커뮤니케이션 방식을 통해 영향력의 크기가 달라지는 양상을 보였다. 따라서 보수의 양이 절대적인 영향력을 행사한다고 보기는 어려우며, 해당 방식이 결과물의 질에 미치는 영향력을 충분히 발휘하기 위해서는 작업자와의 커뮤니케이션을 적절히 활용해야 한다는 점을 알 수 있었다.

#### 6.3.4 단계목표+최종목표와 최종목표의 효과 비교



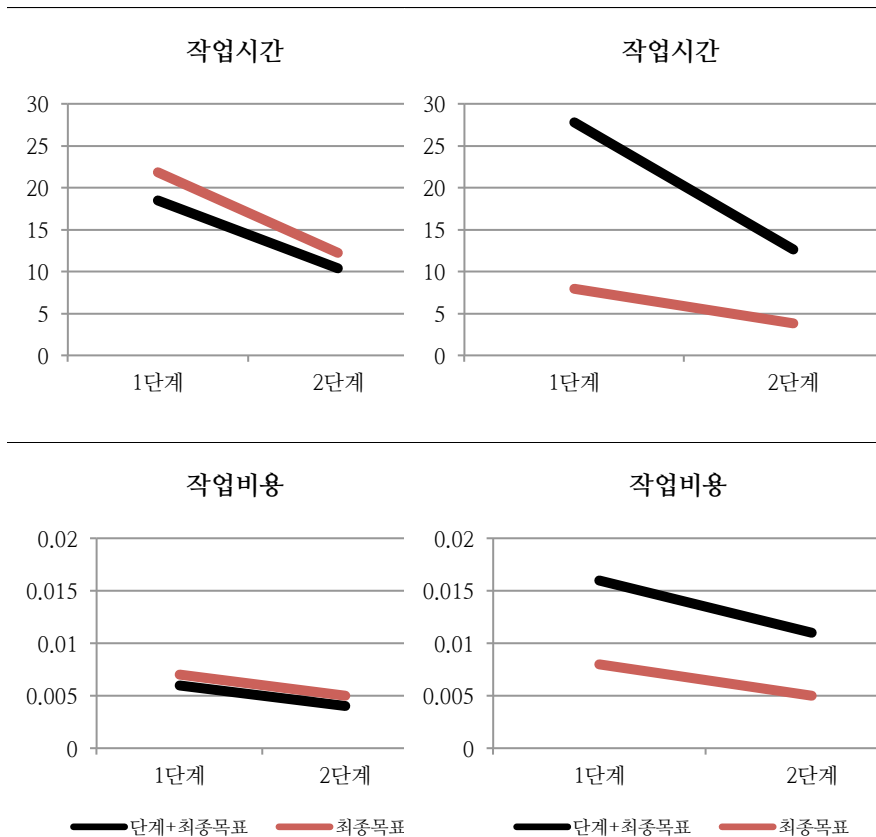


그림 20. 단계목표+최종목표와 최종목표의 단계별 종속변인

여기서부터는 제공된 목표가 결과물의 질에 미치는 영향력을 살펴본다. 먼저 목표의 세부성에 따른 영향력의 차이를 비교하기 위해 단계목표와 최종목표를 함께 제공하는 조건과 최종목표만을 제공하는 조건을 비교하였다. 먼저 각 목표가 참여자의 행동 변화를 이끌어냈는지 살펴보기 위해 1단계와 2단계의 종속변인에 대한 대응표본 t 검정을 실시했다. 그 결과 단계목표와 최종목표를 함께 제공했을 때 \$0.15 지급 조건에서 작업길이는 1단계( $M=59.47$ ,  $SD=23.30$ )보다 2단계( $M=53.37$ ,  $SD=23.27$ )에서

유의하게 감소했다( $t(18)=5.87, p < .05$ ). 작업시간도 1단계( $M=18.47, SD=16.97$ )와 2단계( $M=10.35, SD=12.57$ ) 간에 유의한 차이를 보였고( $t(18)=6.01, p < .05$ ), 작업비용의 경우에도 1단계( $M=0.006, SD=0.005$ )보다 2단계( $M=0.004, SD=0.003$ )에서 유의미하게 적었다( $t(18)=3.08, p < .05$ ). 이처럼 단계목표와 최종목표를 함께 제공했을 때 작업 목적 부합성 및 효율성이 높아지는 결과를 얻을 수 있었다. 이러한 결과는 \$0.3을 지급하는 조건에서도 동일하게 나타났다. 1단계의 작업길이( $M=67.43, SD=23.84$ )보다 2단계의 작업길이( $M=59, SD=24.27$ )가 유의미하게 짧았고( $t(22)=6.56, p < .05$ ), 작업시간이 1단계( $M=27.79, SD=27.29$ )보다 2단계( $M=12.63, SD=14.55$ )에서 감소했으며( $t(22)=4.49, p < .05$ ), 작업비용이 1단계( $M=0.016, SD=0.013$ )보다 2단계( $M=0.011, SD=0.007$ )에서 줄었다( $t(22)=3.76, p < .05$ ). 즉 \$0.15 및 \$0.3을 지급하는 조건 모두에서 단계목표와 최종목표를 함께 제공할 시에 작업길이와 시간 및 비용은 감소하였다.

같은 작업을 최종목표만 제공한 경우에 대해서도 실시한 결과 작업길이와 시간 및 비용은 모든 조건에서 유의미하게 감소하였다. 먼저 \$0.15를 지급하는 조건에서 작업길이는 1단계( $M=67.14, SD=18.88$ )와 2단계( $M=59.5, SD=18.19$ )에서 유의하게 차이가 났고( $t(21)=7.03, p < .05$ ), 작업시간도 1단계( $M=21.86, SD=17.06$ )보다 2단계( $M=12.23, SD=12.43$ )에서 감소하였다 ( $t(21)=5.36, p < .05$ ).

작업비용도 1단계( $M=0.007$ ,  $SD=0.004$ )보다 2단계( $M=0.005$ ,  $SD=0.003$ )에서 유의하게 줄었다( $t(21)=3.93$ ,  $p < .05$ ). 즉, 단계목표와 최종목표를 함께 제공할 때와 마찬가지로 작업의 목적 부합성 및 효율성이 높은 결과물이 도출되었다. \$0.3을 지급하는 경우에도 작업길이는 1단계( $M=56.78$ ,  $SD=13.60$ )보다 2단계( $M=41.61$ ,  $SD=12.63$ )에서 유의하게 줄었고 ( $t(22)=6.62$ ,  $p < .05$ ), 작업시간도 1단계( $M=7.92$ ,  $SD=8.12$ )보다 2단계( $M=3.81$ ,  $SD=4.31$ )에서 유의하게 단축되었으며( $t(22)=3.40$ ,  $p < .05$ ), 작업비용 또한 1단계( $M=0.008$ ,  $SD=0.002$ )보다 2단계( $M=0.005$ ,  $SD=0.002$ )에서 감소했다 ( $t(22)=8.58$ ,  $p < .05$ ). 정리하면 단계목표와 최종목표를 함께 제공하는 경우 \$0.15를 지급하는 경우와 \$0.3을 지급하는 경우 모두에서 작업길이, 시간, 비용이 모두 유의하게 감소하는 양상을 보였다.

다음으로 조건별로 중속변인의 변화폭의 차이를 비교하기 위해 t 검정을 실시하였다. \$0.15를 지급하는 경우에 단계목표+최종목표 조건과 최종목표 조건 간 작업길이, 작업시간 및 작업비용 변화폭의 유의미한 차이는 없었다. 반면 \$0.3을 지급하는 경우에는 두 조건에서 작업길이, 작업시간 및 작업비용 모두에서 변화폭에 유의한 차이가 발견되었다. 작업길이의 변화폭은 최종목표 조건( $M=-15.17$ ,  $SD=10.99$ )이 단계목표+최종목표 조건( $M=-8.43$ ,  $SD=6.16$ )보다 유의하게 컸다( $t(34.58)=2.56$ ,  $p < .05$ ).

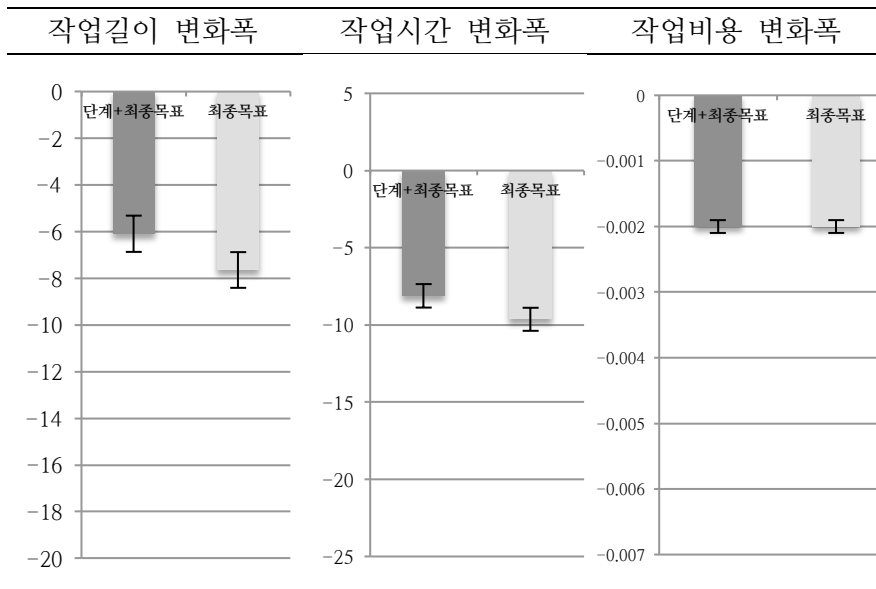
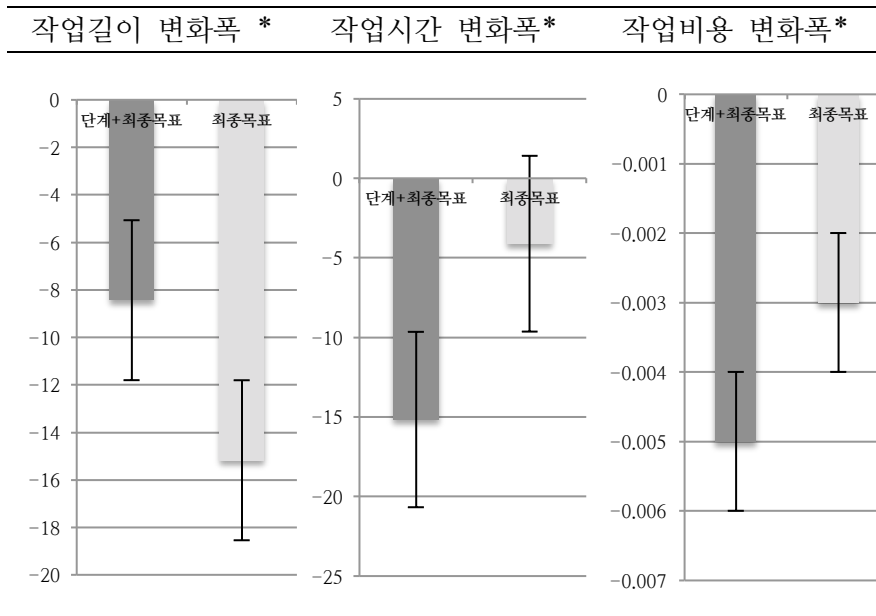


그림 21. 단계목표+최종목표 vs. 최종목표\_종속변인 변화폭(\$0.15)



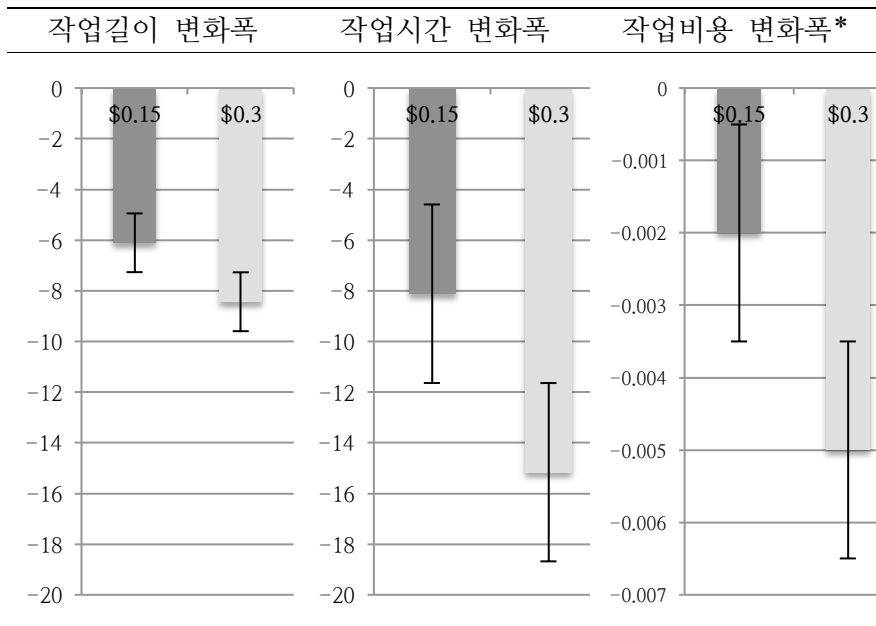
\*  $p < .05$

그림 22. 단계목표+최종목표 vs. 최종목표\_종속변인 변화폭(\$0.3)

이 결과에서 작업의 목적 부합성 측면에서의 결과물의 질은 최종적으로 달성해야 할 지점을 알려주는 경우에 더 향상된다는 점을 알 수 있었다. 그러나 작업시간의 변화폭은 단계목표+최종목표 조건( $M=-14.16$ ,  $SD=15.12$ )이 최종목표 조건( $M=-4.11$ ,  $SD=5.80$ )에서보다 크게 나타났다( $t(28.34)=-2.98$ ,  $p < .05$ ). 작업비용도 작업시간과 마찬가지로 단계목표+최종목표 조건에서의 변화폭( $M=-0.006$ ,  $SD=0.007$ )이 최종목표 조건의 변화폭( $M=-0.002$ ,  $SD=0.001$ )보다 유의미하게 컸다( $t(23.149)=-2.34$ ,  $p < .05$ ). 즉 작업의 효율성 측면에서 결과물의 질을 향상시키는데에는 세부적인 목표가 보다 효과적임을 알 수 있었다.

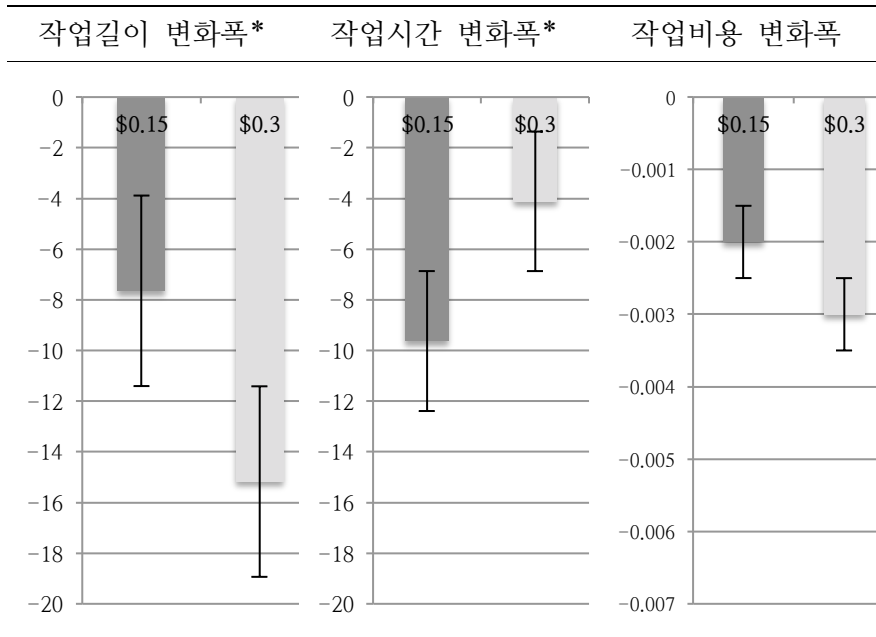
다음 단계로 보수의 차이가 결과물의 질적 향상에 미치는 영향을 비교하기 위해 같은 커뮤니케이션 조건 내에서 보수조건 별로 종속변인의 변화폭을 비교하였다. 먼저 단계목표와 최종목표를 함께 제공하는 조건에서는 작업비용의 변화폭에서 유의미한 차이가 발견되었다. \$0.15를 제공하는 경우에 작업비용의 변화폭( $M=-0.001$ ,  $SD=0.002$ )은 \$0.03을 제공하는 경우의 변화폭( $M=-0.006$ ,  $SD=0.007$ )과 유의한 차이를 보여( $t(26.4)=2.64$ ,  $p < .05$ ), 보수의 양이 늘어났을 때 비용의 효율성이 높아지는 양상을 보였다. 반면 최종목표만을 제공하는 경우에는 보수의 양에 따른 작업길이와 작업시간의 차이가 있었다. \$0.15 지급 조건의 작업길이 변화폭( $M=-7.34$ ,  $SD=5.09$ )과 \$0.3 지급 조건의 길이

변화폭( $M=-15.17$ ,  $SD=11.00$ )은 유의한 차이를 보였다( $t(31.33)=2.97$ ,  $p < .05$ ). 이처럼 최종목표 조건에서는 보수의 양이 증가했을 때 보다 목적에 부합하는 결과물을 도출할 수 있었다. 그러나 작업시간의 경우 \$0.15를 지급하는 조건에서의 변화폭( $M=-9.63$ ,  $SD=8.43$ )이 \$0.3을 지급하는 경우의 변화폭( $M=-1.13$ ,  $SD=5.80$ )보다 유의미하게 컸다( $t(37.10)=-2.55$ ,  $p < .05$ ). 따라서 작업의 목적 부합성 측면과 달리 시간 효율성 측면에서는 보수의 양이 적을 때 보다 결과물의 질이 향상되는 것으로 나타났다.



\*  $p < .05$

그림 23. 단계목표+최종목표\_ 보수조건별 종속변인 변화폭



\*  $p < .05$

그림 24. 최종목표\_ 보수조건별 종속변인 변화폭

위와 같은 결과를 종합해보면 단계목표+최종목표 조건과 최종목표 조건은 \$0.3 지급 조건에서만 종속변인의 변화폭에 유의한 차이가 있었고, \$0.15 지급 시에는 차이가 없었다. 그리고 \$0.3 지급 시에 최종목표 조건에서는 단계목표+최종목표 조건에서보다 작업길이가 보다 큰 폭으로 줄었으나 작업시간과 비용은 오히려 단계목표+최종목표 조건에서 더 크게 감소했다. 가설 2.1에서는 단계목표+최종목표 조건이 최종목표 조건보다 결과물의 질을 높일 것으로 보았는데, 실험 결과 가설 2.1의 일부만을 지지할 수 있었다. 작업길이를 줄이는 작업의 목적 달성 측면에서는 최종목표 조건이 더 효과적이었으나, 작업시간과

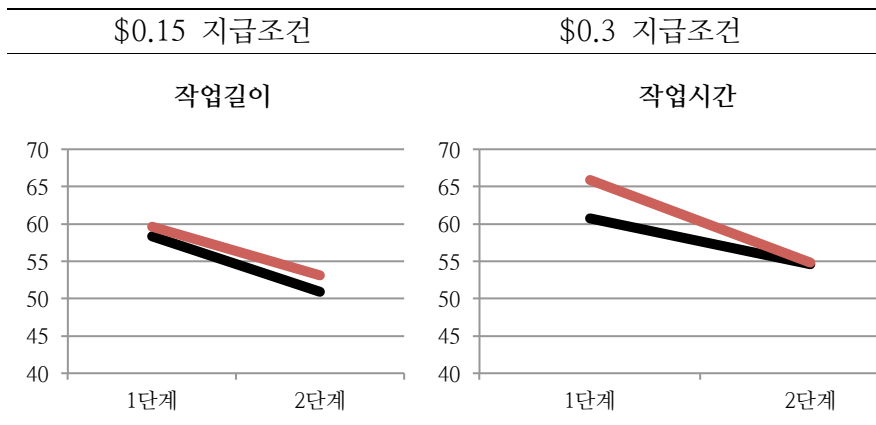


비용을 줄이는 작업의 효율 측면에서는 단계목표+최종목표 조건이 더 효과적이었기 때문이다. 그러나 단계목표+최종목표 조건의 작업길이 감소폭이 최종목표 조건에 비해 상대적으로 작을 뿐이며, 긍정적 피드백이나 기타 다른 커뮤니케이션 조건에서의 감소폭보다 크거나 유사하다. 따라서 작업 목적을 달성하는 측면에서 결과물의 질을 충분히 높이고 있었으므로 단계목표와 최종목표를 함께 제공하는 방식은 최종목표만을 제공하는 방식보다 결과물의 질을 더욱 향상시키는 것으로 볼 수 있다.

또한 보수의 양이 결과물의 질적 향상에 미치는 영향력을 알 수 있었는데, 서술적 피드백과 평가적 피드백을 비교할 때와 마찬가지로 영향력의 크기와 방향이 커뮤니케이션 조건에 따라 다르게 나타났다. 절대적인 수치를 살펴보면 단계목표와 최종목표를 함께 제공할 때는 \$0.3을 지급할 때 작업길이, 시간 및 비용이 \$0.15를 지급할 때보다 더 많이 감소했다. 그 중 작업비용의 변화폭이 \$0.3 지급 시에 유의미하게 커 보수의 양이 많을 때 결과물의 비용 효율 측면이 향상되는 것으로 나타났다. 반면 최종목표만 제공할 때 작업길이와 비용의 절대적인 감소폭은 \$0.3 지급 시에 더 컸지만 작업시간은 \$0.15 지급 시에 더 컸다. 그리고 그 중에서 작업길이를 더 큰 폭으로 감소시키는 조건은 \$0.3을 지급할 때지만, 작업시간을 더 많이 줄인 조건은 \$0.15 지급 시로 나타났다.

위와 같은 결과에서 보수의 증가는 결과물의 질적 향상에 긍정적인 영향을 미친다는 점을 알 수 있었다. 그러나 작업 목적 부합성 및 시간 효율성의 향상 효과는 최종목표에서만 나타나고, 비용의 효율성을 높이는 효과는 단계목표와 최종목표 조건에서만 나타나는 등 보수의 양이 갖는 영향력은 커뮤니케이션 조건에 따라 다르게 나타났다. 목표가 세부적일 때는 보수가 증가했을 때 비용의 효율성이 높아진 반면 목표가 단순할 때는 목적 부합성과 시간 효율성이 높아졌다. 이러한 결과를 통해 볼 때, 보수의 양이 갖는 영향력을 보다 적절히 활용하기 위해서는 작업자와의 커뮤니케이션을 고려하는 것이 바람직할 것이다. 예를 들어 본 실험 결과에 근거하여 클라우드소싱 결과물의 질을 높이기 위해 세부적인 목표를 제시하는 경우, 평균 수준보다 높은 보수를 지급함으로써 비용의 효율성을 도모할 수 있을 것이다.

### 6.3.5 표적형 목표와 성취형 목표의 효과 비교



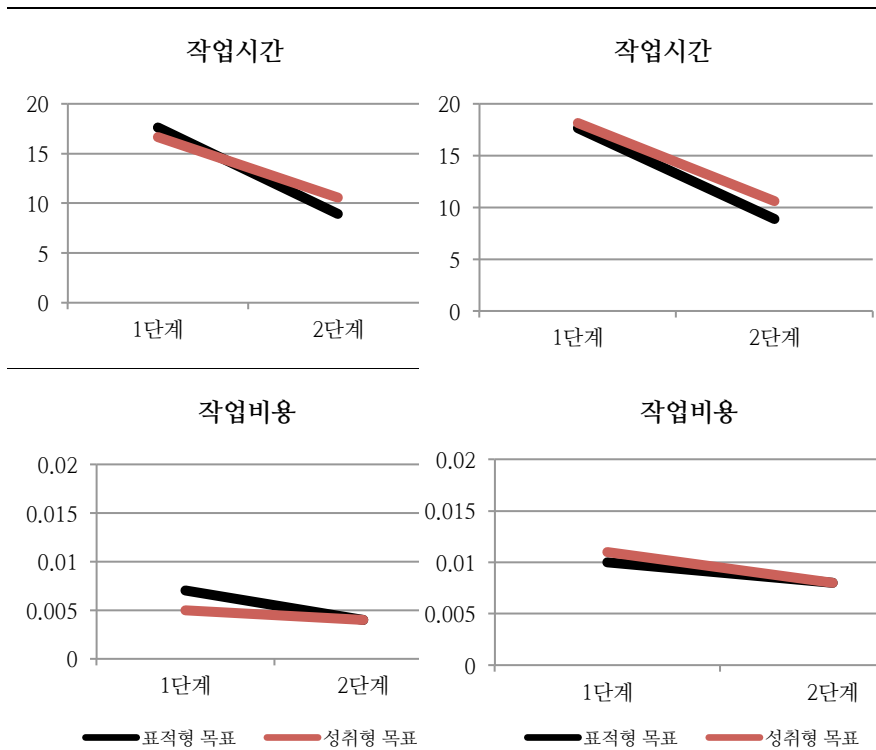


그림 25. 표적형 목표와 성취형 목표의 단계별 종속변인

표적형 목표와 성취형 목표가 참여자의 행동 변화를 이끌어내는지 알아보기 위해 대응표본 t 검정을 실시하였다. 먼저 표적형 목표를 제공했을 때 \$0.15를 지급한 경우 작업길이, 시간 및 비용이 모두 감소하였다. 1단계의 작업길이( $M=58.32$ ,  $SD=23.86$ )는 2단계( $M=50.91$ ,  $SD=23.14$ )와 유의한 차이를 보였고( $t(21)=4.73$ ,  $p < .05$ ), 작업시간은 1단계( $M=17.61$ ,  $SD=21.27$ )보다 2단계( $M=8.93$ ,  $SD=12.07$ )에서 유의하게 감소했다( $t(21)=3.73$ ,  $p < .05$ ). 작업비용도 1단계( $M=0.007$ ,  $SD=0.008$ )에 비해 2단계( $M=0.004$ ,  $SD=0.004$ )에서 유의미하게

줄었다( $t(21)=2.55, p < .05$ ). 표적형 목표를 제시했을 때 모든 측면에서 결과물의 질이 향상된 것인데, 이러한 결과는 \$0.3 지급 조건에서도 동일했다. 1단계의 작업길이( $M=60.7, SD=19.80$ )와 2단계 작업길이( $M=54.6, SD=18.78$ )는 유의한 차이를 나타냈고( $t(19)=4.32, p < .05$ ), 작업시간도 1단계( $M=17.68, SD=11.56$ )보다 2단계( $M=8.88, SD=5.88$ )에서 감소했다( $t(19)=4.69, p < .05$ ). 또 1단계 작업비용( $M=0.010, SD=0.006$ )보다 2단계 작업비용( $M=0.008, SD=0.005$ )이 유의하게 줄었다( $t(19)=5.44, p < .05$ ). 이와 같이 볼 때 표적형 목표는 참여자의 작업 변화를 이끌어내며, 두 가지 보수 조건 모두에서 작업의 목적 부합성과 시간 및 비용의 효율성을 높이는 결과를 가져왔다.

이어서 성취형 목표에 대해서도 같은 분석을 실시했다. \$0.15 지급 조건에서 살펴보면 1단계 작업길이( $M=59.59, SD=22.39$ )와 2단계 작업길이( $M=53.09, SD=21.86$ )는 유의한 차이를 보였다( $t(21)=3.26, p < .05$ ). 작업시간도 1단계의 작업시간( $M=16.66, SD=12.58$ )보다 2단계의 작업시간( $M=10.57, SD=9.80$ )이 유의하게 짧았고( $t(21)=4.31, p < .05$ ), 작업비용에서도 1단계( $M=0.005, SD=0.004$ )보다 2단계( $M=0.004, SD=0.003$ )에서 유의하게 적었다( $t(21)=3.23, p < .05$ ). 그러므로 세 가지 측면에서 모두 결과물의 질이 향상되었다고 볼 수 있었다. 아울러 \$0.3 지급 조건에서도 1단계 작업길이( $M=65.85, SD=15.92$ )보다 2단계

작업길이( $M=54.8$ ,  $SD=19.70$ )가 유의하게 짧았고( $t(19)=7.70$ ,  $p < .05$ ), 1단계 작업시간( $M=18.12$ ,  $SD=11.35$ )보다 2단계 작업시간( $M=10.58$ ,  $SD=6.95$ )이 줄었다( $t(19)=4.63$ ,  $p < .05$ ). 아울러 작업비용도 차이를 보였는데, 1단계에 필요한 비용( $M=0.011$ ,  $SD=0.006$ )보다 2단계에 소요된 비용( $M=0.008$ ,  $SD=0.004$ )이 유의하게 감소했다( $t(19)=6.08$ ,  $p < .05$ ). 이와 같이 성취형 목표도 두 가지 보수조건에서 결과물의 질을 향상시킨다는 점을 알 수 있었다.

위와 같이 참여자의 행동 변화를 이끌어내는 점을 확인한 후, 그 변화의 폭을 비교하기 위해  $t$  검정을 실시하였다. \$0.15 지급 조건에서는 작업길이, 작업시간, 작업비용 모두에서 표적형 목표와 성취형 목표 조건의 변화폭 간에 유의미한 차이가 발견되지 않았다. 반면 \$0.3 지급 조건에서는 성취형 목표가 표적형 목표보다 작업길이와 작업비용을 많이 감소시켰고, 작업시간에는 유의한 차이를 나타내지 않았다. 표적형 목표 조건의 작업길이 감소폭( $M=-6.10$ ,  $SD=6.32$ )과 성취형 목표 조건의 감소폭( $M=-11.05$ ,  $SD=6.42$ )은 유의미한 차이를 나타내( $t(37.99)=2.46$ ,  $p < .05$ ), 성취형 목표가 보다 효과적으로 작업 목적에 부합하는 결과물을 도출한다는 점을 알 수 있었다. 또한 작업비용에서도 표적형 목표 조건의 비용 감소폭( $M=-0.002$ ,  $SD=0.001$ )은 성취형 목표 조건의 비용 감소폭( $M=-0.003$ ,  $SD=0.002$ )과 유의한 차이를

보였다( $t(31.07)=2.42, p < .05$ ). 작업의 목적 부합성 측면과 마찬가지로 작업의 비용 효율성을 높이는데에 성취형 목표는 표적형 목표보다 효과적이었다. 즉, 성취형 목표가 표적형 목표보다 작업 목적 부합성 및 비용 효율성 등 결과물의 질을 나타내는 3가지 측면 중 2가지에서 더 효과적인 것으로 나타났다.

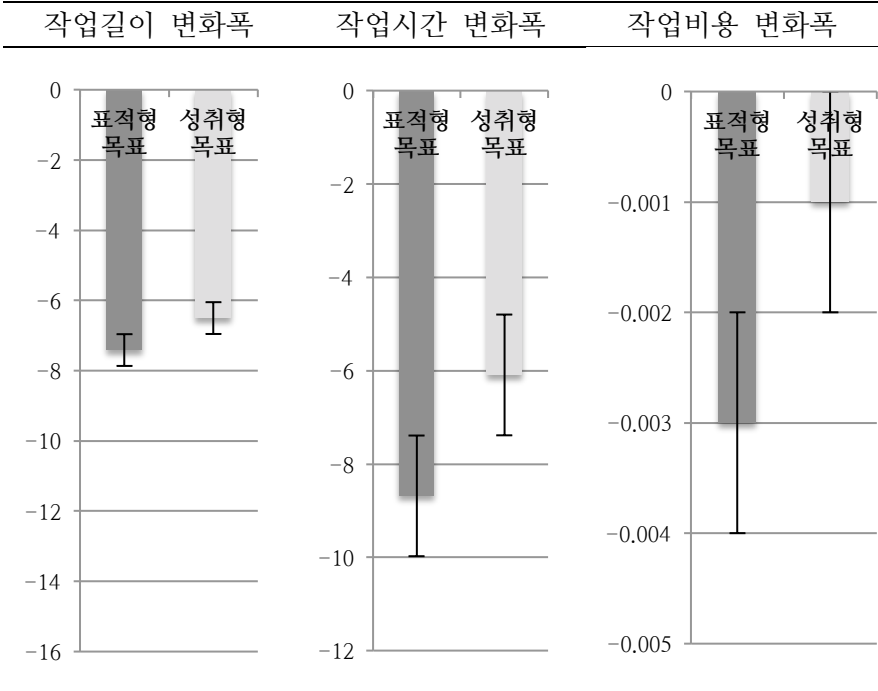
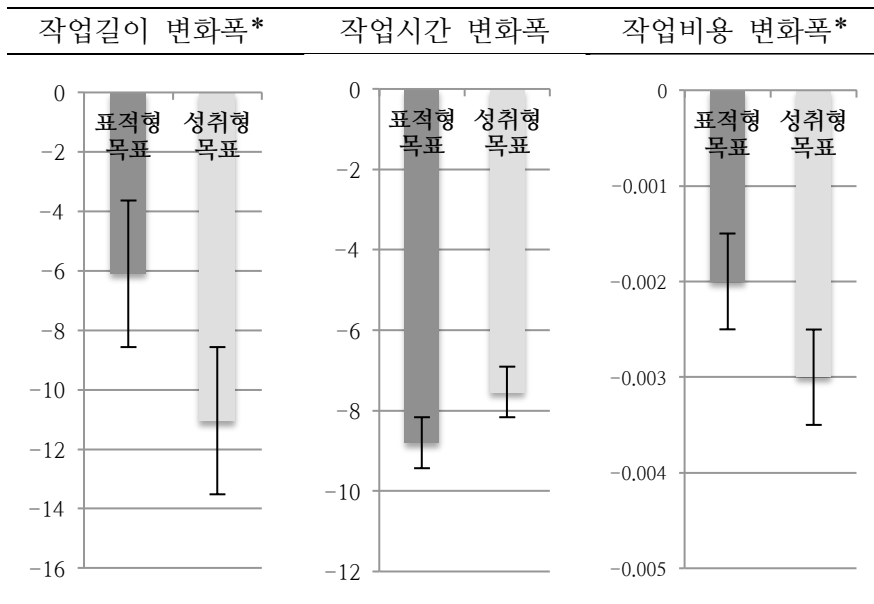


그림 26. 표적형 목표 vs. 성취형 목표\_ 종속변인 변화폭(\$0.15)



\*  $p < .05$

그림 27. 표적형 목표 vs. 성취형 목표\_ 종속변인 변화폭(\$0.3)

이어서 보수의 양에 따른 차이를 살펴보기 위해 동일한 목표 조건에서 비용의 차이에 따른 종속변인의 변화폭을 비교하였다. 이때 표적형 목표를 제시하는 경우에는 작업길이, 작업시간, 작업비용 등 3가지 종속변인 모두에서 유의미한 차이가 발견되지 않았다. 절대적인 수치를 살펴보아도 \$0.15 지급 시에 작업길이와 작업비용은 \$0.3을 지급할 때보다 오히려 더 많이 줄어들었고, 작업시간에는 큰 차이가 보이지 않았다. 그러므로 표적형 목표를 제시할 때는 보수의 양이 큰 영향력을 미치지 못함을 알 수 있었다. 반면 성취형 목표를 제시할 때는 \$0.15 지급 시 작업비용의 감소폭( $M=-0.001$ ,  $SD=0.002$ )보다 \$0.3 지급 시 작업비용

감소폭( $M=-0.003$ ,  $SD=0.002$ )이 유의하게 컸다( $t(37.78)=2.64$ ,  $p < .05$ ). 또한 유의미한 차이는 나타나지 않았지만 \$0.3 지급 시의 작업길이 및 시간이 줄어든 정도는 \$0.15 지급 시보다 컸다. 그러므로 성취형 목표가 결과물의 질에 미치는 영향력은 보수의 양에 영향을 받으며, 보수의 양이 증가할 때 결과물의 질이 보다 향상될 수 있음을 알 수 있었다.

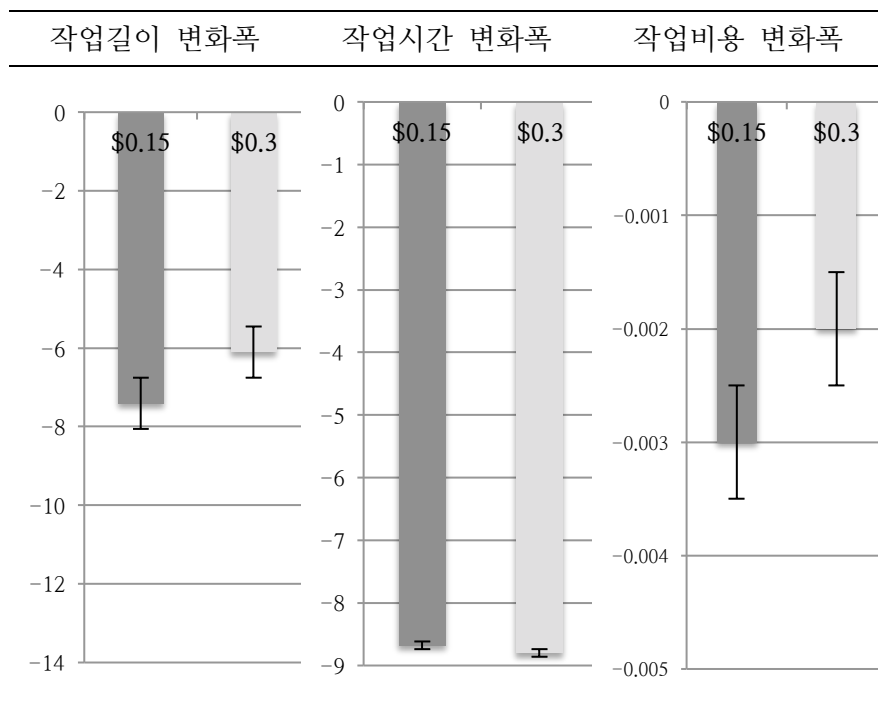
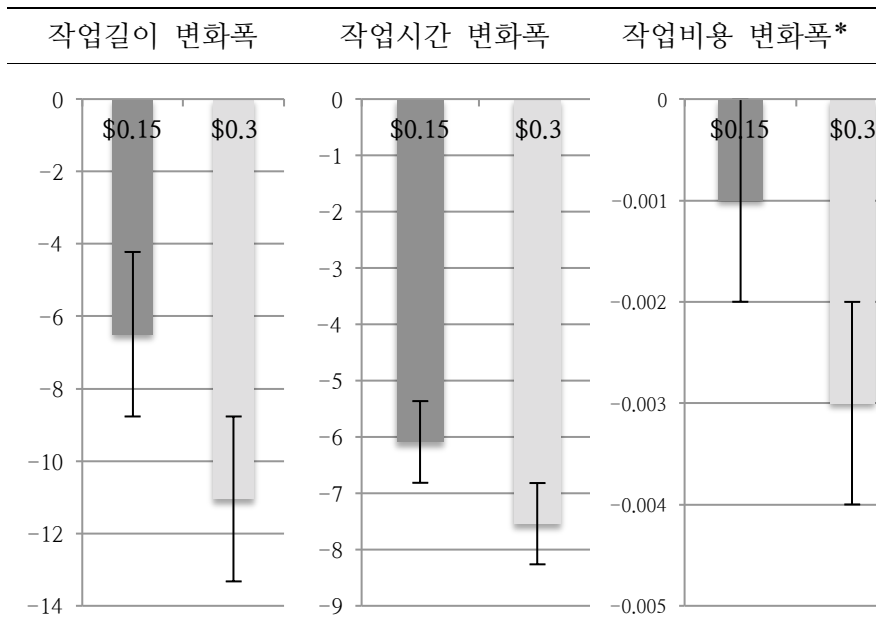


그림 28. 표적형 목표\_ 보수조건별 종속변인 변화폭





\*  $p < .05$

그림 29. 성취형 목표\_ 보수조건별 종속변인 변화폭

위와 같은 결과에서 성취형 목표는 표적형 목표보다 결과물의 질적 향상에 효과적이라는 가설 2.2을 지지할 수 있었다. 성취형 목표는 \$0.3 지급 조건에서 표적형 목표에 비해 작업길이 및 비용을 더 크게 감소시켜 보다 효과적으로 결과물의 질을 향상시키고 있었다. \$0.15 지급 조건에서 종속변인의 감소폭 자체는 표적형 목표 조건에서 더 크게 나타났지만 그 차이가 유의미하지 않아 표적형 목표가 성취형 목표보다 결과물의 질적 향상에 효과적이라고 보기는 어려웠다. 이러한 차이가 나타난 이유에는 여러가지가 있을 수 있는데, 그 중 작업자가 작업의

이유를 알고 나서 작업 자체에 대한 이해도나 관여도가 높아졌을 가능성을 고려해볼 수 있다.

본 실험과 같이 단순작업이 아닌 경우에는 작업을 해야 하는 배경을 설명해줄 때, 의뢰자가 요구하는 사항들을 보다 잘 이해할 수 있었을 것이다. 예를 들어 최종 보고서의 글자 수 제한을 맞추기 위해 길이를 줄인다는 배경을 설명해준다면, 제시문의 본 의도를 훼손하지 말라는 요구사항이 제시되었는지 이해할 수 있다. 이렇게 된다면 가이드 라인을 단순한 제한사항으로 받아들이지 않고, 보다 작업에 집중할 수 있을 것이다. 그러므로 비대면 근무 환경이고 상대적으로 단순한 작업이 이루어지는 클라우드소싱에서도 작업자가 작업을 충분히 이해하고 공감할 수 있도록 커뮤니케이션하는 것이 중요하다고 할 수 있다.

아울러 보수의 양은 앞선 분석과 마찬가지로 결과물의 질을 높이는데 긍정적인 영향력을 미치고 있었다. 성취형 목표가 표적형 목표보다 결과물의 질을 효과적으로 향상시킨 결과는 \$0.3 지급 조건에서만 나타나았고, \$0.15 지급 시에는 유의미한 차이가 나타나지 않았다. 그러므로 작업의 보수 수준이 높을 때는 성취형 목표를 제시하는 것이 표적형 목표를 제시하는 방식보다 효과적이나, 보수 수준이 낮을 때는 성취형 목표의 영향력이 뚜렷하게 나타나지 않을 수 있다는 점을 발견하였다. 그러나 보수의 양을 늘리는 방식이 절대적으로 결과물의 질적 향상을

보장해주지는 못했다. 성취형 목표의 경우 각 변인의 절대값을 보면 보수의 수준이 높아질 때 종속변인의 변화폭이 커지는 것처럼 보였으나 유의미한 차이는 작업비용에만 나타났다. 또한 표적형 목표를 제공할 때는 유의미한 차이가 나타나지 않았고, 절대값을 보더라도 \$0.15를 지급할 때 작업길이와 비용의 감소폭이 더 크게 나타나기도 했다.

이렇게 볼 때 보수의 양적 증가가 결과물의 질에 미치는 긍정적인 영향력은 커뮤니케이션 조건에 따라 다르게 나타나고, 반대로 특정 커뮤니케이션이 결과물의 질에 미치는 영향력은 보수 수준에 따라 다르게 나타날 수 있다. 그러므로 작업자와의 커뮤니케이션 방식을 고려하여 보수를 지급할 때 가장 효과적으로 결과물의 질을 높일 수 있다는 점을 알 수 있었다.

## 6.4 연구결과 요약

실험 결과 크라우드소싱에서 결과물의 질을 높이는데 효과적인 4가지의 커뮤니케이션 방식을 발견하였고, 보수의 양이 결과물의 질적 향상에 미치는 영향력을 살펴볼 수 있었다. 먼저 부정적 피드백이 긍정적 피드백보다 작업의 목적에 부합하는 결과물을 효율적으로 도출하였다. 또 평가적 피드백은 서술적 피드백보다 작업의 목적에 부합하고 시간과 비용을 절약하는 등 전반적으로 결과물의 질을 높이는 것으로 나타났다. 그 다음 단계목표와

최종목표를 함께 제공했을 때는 최종목표만 제시했을 때보다 효율적으로 결과물을 도출할 수 있었다. 마지막으로 성취형 목표는 표적형 목표보다 작업 목적에 부합하고 비용 효율성이 높은 결과물을 도출하였다. 아울러 보수의 양은 결과물의 질에 긍정적인 영향력을 끼쳤으나, 구체적인 크기 및 방향은 커뮤니케이션 조건에 따라 다르게 나타나는 등 절대적이지 못한 것으로 나타났다.

## 7. 논의 및 시사점

지금까지 본 연구는 클라우드소싱에서 결과물의 질을 높이기 위한 방안에 대해 살펴보았다. 기존의 연구가 결과물의 질을 높이기 위한 방안으로 보수의 양을 늘리거나 필터링을 통해 결과물의 질을 관리하는 방법에 초점을 맞추었던 것에 반해, 본 연구는 의뢰자와 작업자 사이의 커뮤니케이션 방법에 초점을 맞추어 진행되었다. 연구를 진행한 결과 커뮤니케이션과 보수 수준 모두 결과물의 질에 일정한 영향력을 행사하고 있었고, 두 방식 중 하나가 결과물의 질에 절대적인 영향력을 행사하지는 않았다. 먼저 작업자와의 커뮤니케이션은 방식에 관계 없이 결과물의 질을 높이는 영향력을 가지고 있었다. 모든 커뮤니케이션 조건에서 1단계와 2단계의 작업길이, 작업시간 및 작업비용은 감소했고, 이는 \$0.15 지급 조건과 \$0.3 지급 조건에서 동일하게 나타났다. 따라서 본 연구가 설정한 3가지 측면에서의 결과물의 질이 커뮤니케이션 메시지를

받은 후 향상된다는 점을 알 수 있었다. 그리고 8개의 커뮤니케이션 조건 중, 작업자에게 명확한 판단 근거를 제시하거나 작업을 이해하는데 더 많은 도움을 줄 수 있는 커뮤니케이션이 결과물의 질적 향상에 효과적이었다. 연구 결과 피드백 중에서는 부정적 피드백과 평가적 피드백이 적합했고, 목표 중에서는 단계+최종목표와 성취형 목표가 적합한 것으로 나타났다.

부정적 피드백이 긍정적 피드백보다 결과물의 질적 향상에 효과적인 이유로 본 연구는 부정적 피드백이 본인의 작업 결과에 대해 다시 한번 생각하게 유도하는 효과를 가지고 있기 때문이라고 보았다. 아울러 크라우드소싱이 비대면 작업환경이고 작업자와 의뢰자가 단기적인 관계를 맺는다는 점도 이러한 결과가 나오는데 일조했을 것으로 보인다. 장기적인 관계를 맺고 대면 접촉하는 경우에는 긍정적인 피드백을 통해 작업자의 기분을 좋게 하고 작업 의욕을 높일 수 있다. 하지만 크라우드소싱은 정반대의 환경이므로 현재 결과물에서 부족한 점을 명확하게 짚어주는 방식은 작업자의 동기에 악영향을 끼치지 않을 것이다. 또 일반 작업환경과 달리 지시를 받거나 도움을 구할 수 없기 때문에 부정적 피드백에서 제공하는 정보는 작업자에게 꼭 필요한 정보였을 가능성이 높다.

평가적 피드백이 서술적 피드백보다 효과적으로 결과물의 질을 높인 결과도 부정적 피드백의 효과와 유사하게 해석할 수 있다. 결과물의 잘하고 못함을 판단하는 것은 대면 작업환경에서

작업자의 자존심을 해치는 등 동기 부여에 악영향을 줄 수 있지만, 비대면 환경에서는 오히려 스스로의 작업 결과를 판단하는데 도움을 줄 수 있다. 클라우드소싱은 철저히 제출된 결과물에 따라 보수가 지급되므로 결과물이 의뢰자가 원하는 수준보다 못하다면 보수를 지급받지 못할 확률이 커진다. 따라서 평가적 피드백은 보수를 받을 수 있는 확률을 높여주고, 작업자에게 보다 명확한 지시를 내린 것이 되므로 결과물의 질을 높이는데 기여했을 것이다.

또한 목표와 관련된 커뮤니케이션에서는 목표가 세부적일 때 결과물의 질이 상대적으로 향상된다는 점을 발견했다. 결과물의 질을 측정하는 3가지 측면 중 작업길이 즉, 작업의 목적 부합성은 최종목표와 같이 포괄적인 목표에서 더 향상되었지만 작업시간과 작업비용 등 작업의 효율 측면은 세부적일 때 향상되었다. 그리고 작업길이의 감소폭도 상대적으로 단계목표+최종목표 조건이 최종목표 조건에 비해 적을 뿐 절대적인 감소폭은 상당한 수준으로 나타났으므로, 종합해볼 때 목표의 세부성이 작업 결과물의 질적 향상에 도움을 준다는 사실을 알 수 있었다. 이러한 결과는 부정적 및 평가적 피드백 제시 결과와 연결해서 이해할 수 있다. 클라우드소싱의 작업자들은 하나의 작업에 오랜 시간을 투자하는 것이 아니라 짧은 시간 동안 여러 개의 서로 다른 작업을 수행한다. 또 의뢰자와는 이메일 등 비동시적인 커뮤니케이션이 주로 이루어져서 의문점을 즉시 해결하기 어려우므로, 현재 작업에 대한

명확한 기준이 필요하다. 목표가 세부적으로 제시되면 보다 정확한 작업을 할 수 있고, 우왕좌왕하는 일이 줄어들기 때문에 작업의 효율이 높아질 수 있다. 따라서 클라우드소싱 환경에서 결과물의 질을 향상시키려면 세부적인 목표가 필요하다.

표적형 목표보다 성취형 목표가 결과물의 질을 향상시키는데 효과적이었다는 점도 주목할 만 하다. 기존에 결과물의 질을 높이기 위해 보수를 많이 지급한 방법은 클라우드소싱 작업자들이 경제적 수입에 관심이 높고, 작업에 대해 기여도나 관심이 낮다는 생각에 기반해있다. 따라서 실제 수행하는 작업의 내용과는 직접적인 관계가 없는 보수의 양을 통해 결과물의 질적 향상을 도모했다. 그러나 본 연구 결과는 클라우드소싱 작업자들도 작업을 해야 하는 이유를 이해하게 되면 보수의 양과 큰 관련이 없이도 양질의 결과물을 도출할 수 있다는 점을 시사하고 있다. 따라서 작업의 배경 및 이유에 대한 설명을 통해 작업자들이 작업 자체에 대해 관여도와 관심을 갖게 하는 방법이 결과물의 질적 향상에 효과적임을 알 수 있었다.

위의 결과와 같이 작업자와의 커뮤니케이션은 결과물의 질을 향상시키는 영향력을 가지고 있었는데, 본 연구는 이에 더해 보수의 양적 증가 방식과 커뮤니케이션의 영향력을 함께 살펴보았다. 양질의 결과물을 얻기 위해 가장 대표적으로 이용되는 방식이 보수의 양을 늘리는 것인만큼, 본 연구의 결과에서도

보수의 수준을 높일 때 결과물의 질에 긍정적인 변화가 관찰되었다. 그러나 그 영향력이 절대적이지 않았고 커뮤니케이션 방식에 따라 영향력의 방향과 크기가 달라져, 결과물의 질을 관리하기 위해서는 지급하는 보수의 양과 함께 작업자와의 커뮤니케이션을 고려해야 할 필요가 있었다. 먼저 긍정적 피드백과 부정적 피드백 조건을 비교했을 때, 보수의 양이 다르더라도 결과물의 질적 향상에는 유의미한 차이가 발견되지 않았다. 또한 서술적 피드백과 평가적 피드백을 비교했을 때는 2가지 보수조건에서 종속변인의 감소 양상이 동일한 방향으로 나타났지만, 피드백 조건의 감소폭 간 차이는 다르게 나타났다. 즉, 보수 수준이 갖는 영향력의 크기는 커뮤니케이션 조건에 따라 달랐다.

반면 목표를 제시하는 경우에는 피드백을 제시할 때보다 보수의 양이 결과물의 질적 향상에 더 많이 기여하는 것으로 나타났다. 단계목표와 최종목표를 함께 제공해 목표가 세부적인 경우에는 최종목표를 제시한 경우보다 시간 및 비용 효율성 측면에서 우수한 결과를 도출했으나 이는 \$0.3 지급 조건에서만 발견되었다. 또 성취형 목표는 표적형 목표보다 작업 목적 부합성 및 비용 효율성 측면에서 효과적이었으나 이또한 \$0.3 지급 시에서 나타난 효과였다. 그러므로 보수의 양이 결과물의 질에 영향력을 미친다고 볼 수 있지만, 반대로 본다면 보수의 양이 갖는 영향력은 커뮤니케이션 조건에 따라 다르게 나타난다고 볼 수 도 있다. 보수



수준을 높였을 때 결과물의 질에 긍정적인 영향을 주는 결과는 목표를 제시하는 방식의 커뮤니케이션 조건에서 주로 발견되었고, 목표 관련 커뮤니케이션 안에서도 보수의 영향력이 일정하지 않게 나타났기 때문이다.

아울러 실험 진행 과정에서 답안의 탈락률이 보수조건에 따라 큰 차이를 보인 점은 흥미로운 부분이었다. 데이터클리닝 과정에서 제외한 답안으로는 내용이 충분하지 못한 경우와 지나치게 성의가 없는 경우가 있었다. 이러한 탈락률을 보수조건별로 비교한 결과 \$0.15 지급 시에는 총 168건 중 3건으로 탈락률은 1.82%였으나, \$0.3 지급 시에는 총 202건 중 23건으로 11.39%에 달해 \$0.15 지급 시 탈락률의 6배가 넘었다. 절대적인 수치나 비율에서 \$0.15 지급 시보다 \$0.3 지급 시에 탈락률이 높았다는 사실은 보수의 양이 많을 때 작업자들이 성실하게 양질의 결과물을 도출할 것이라는 기존의 주장과 상반된다. 각 실험조건에서 보수조건 별로 종속변인의 차이나 변화폭을 검증했을 때는 보수의 양이 결과물의 질이 향상되는지 큰 차이를 가져오지 않는다는 점만을 알 수 있었다. 하지만 이와 같은 탈락률의 차이를 통해 보수의 양은 결과물의 질에 큰 차이를 가져오지 않을 뿐만 아니라, 작업자들의 성실한 참여를 유도하는데 오히려 악영향을 줄 수도 있다는 점이 밝혀졌다. 동일한 텍스트 길이 단축 작업에 참여했을 때 적은 보수에도 끝까지 참여한 사람들은 작업 자체에 대한 관심이

상대적으로 더 많았을 가능성이 있다. 본 실험에서는 낮은 보수 조건에 해당하는 경우 제시문의 1%를 줄이는데 \$0.15를 지급했고, 높은 보수 조건에 해당하는 경우에는 1%를 줄일 때 \$0.3을 지급했다. 물론 낮은 보수조건인 금액도 실질적으로는 크라우드소싱에서 상당히 큰 금액이지만, 작업의 내용이 그만큼 까다로웠으므로 결국 작업을 흥미롭게 느끼고 참여하고 싶어하는 작업자들이 해당 조건에서 작업을 마쳤을 것이다. 반면 1% 당 \$0.3을 제시했을 때는 크라우드소싱 평균 보수와의 차이가 \$0.15 지급 시보다 훨씬 커지기 때문에, 작업길이의 1%를 더 줄였을 때 얻는 이익에 더욱 집중했을 수 있다. 이 경우 작업의 질 보다는 제시문을 극단적으로 줄여 더 많은 보수를 받고자 하는 동기가 나타날 수 있으며, 이것이 높은 탈락률의 원인이 될 수 있다.

위와 같은 연구결과를 종합해보면 보수의 양적 증가가 결과물의 질적 향상에 미치는 영향력을 무시할 수는 없으나 절대적인 영향력을 가지고 있지 않았고, 그 방향과 크기가 작업자와의 커뮤니케이션에 따라 달라졌다. 또한 보수 수준이 높을 때 답안 탈락률이 높아지는 등 부정적인 영향력도 발견되었다. 그러므로 보수 수준과 작업자와의 커뮤니케이션이 크라우드소싱 결과물의 질을 높이는 영향력을 최대한 활용하기 위해서는 두 가지 방식을 함께 고려해야 한다는 결론을 내릴 수 있었다. 보수의 양이 가지는 영향력을 부정하지 않지만 보수의 양에만 집중하지 않았다는

점에서 본 연구는 클라우드소싱 결과물의 질적 향상에 대한 기존 연구들의 한계를 보완할 수 있을 것이다.

본 연구에서는 이처럼 흥미로운 결과들이 도출되었으나, 더욱 좋은 결과를 내기 위해 후속연구에서 극복해야 할 한계점도 발견되었다. 본 연구의 한계로는 먼저 실험 내용이 일반적으로 클라우드소싱에서 진행되는 작업들과 다소 차이가 난다는 점이 있다. 본 연구는 클라우드소싱의 가능성이 점차 확대됨에 따라 높은 난이도의 과업을 수행할 때 결과물의 질을 보장하기 위해서는 기존과 다른 방식이 필요하다는 생각에서 출발했다. 그러나 아직까지 사용자들은 본 연구에서 시행한 작업보다 단순반복적인 작업에 익숙해 이러한 과업을 어려워하고 있었다. 실제로 실험 사이트에 등록된 작업자 중 절반 이상이 중도 포기하는 현상도 나타났다. 따라서 현 연구결과를 즉각적으로 클라우드소싱 환경의 디자인에 적용하기에는 무리가 있을 것이다. 그러나 앞으로 클라우드소싱은 단순작업부터 복잡한 작업에 이르기까지 다양하게 활용되는 방향으로 발전해 나갈 것이므로, 본 연구의 결과를 적용하여 결과물의 질을 안정적으로 관리할 수 있을 것이다.

또한 본 연구에서 작업자와 커뮤니케이션하기 위해 사용된 메시지들은 미리 제작된 것이어서 작업자의 실제 결과물에 대한 평가 내용을 담지 못했다. 본 연구는 비대면 온라인 실험을 대규모로 진행했고, 실험 참여자가 미국부터 인도까지 다양한

국가에 분포되어 있어 누구나 쉽게 이해할 수 있어야 하므로 메시지의 표현 방식이 제한될 수 밖에 없었다. 하지만 이처럼 메시지의 내용이 제한될 경우 커뮤니케이션이 결과물의 질적 향상에 미치는 영향력 또한 제한될 수 밖에 없다. 그러므로 차후에 참여자의 국적이나 특성을 파악하여 메시지를 제공할 수 있다면, 보다 풍부한 커뮤니케이션을 통해 결과물의 질적 향상 폭을 증대할 수 있을 것이다.

## 8. 결론

본 연구는 클라우드소싱이라는 새로운 작업환경에서 결과물의 질을 향상시키기 위한 방안으로 커뮤니케이션에 주목하고, 커뮤니케이션의 방식 및 내용에 따른 영향력의 차이를 비교 검증하였다. 이때 커뮤니케이션 방식은 피드백 제시와 목표 제시 등 2가지로 나뉘고, 긍정적 피드백과 부정적 피드백, 서술적 피드백과 평가적 피드백을 서로 비교했다. 목표의 경우 단계목표+최종목표와 최종목표, 표적형 목표와 성취형 목표의 영향력을 비교했다. 아울러 기존에 시도된 방법 중 보수의 양을 달리하는 방식과 커뮤니케이션 제공 방식의 영향력을 비교하기 위해 두 가지 보수 조건을 적용하였다. 이와 같은 조건을 설정하여 실험을 진행한 결과 클라우드소싱에서 결과물의 질을 효과적으로 향상시킨 커뮤니케이션은 부정적 피드백, 평가적 피드백,

단계목표+최종목표, 그리고 성취형 목표로 밝혀졌다. 이러한 커뮤니케이션 방식들은 비대면 작업환경에서 일하는 클라우드소싱 작업자에게 현재 진행하는 작업을 진행하는데 필요한 즉각적이고 명확한 판단 기준을 제공해주고, 작업에 대한 이해도를 높여 결과물의 질을 높이고 있었다. 또한 보수의 수준이 결과물의 질에 일정한 영향력을 가지고 있었으나 이것이 절대적이지 못해 보수의 양적 증가 방식에 대해 새롭게 바라볼 수 있었다. 보수 수준이 결과물의 질적 향상에 미치는 영향력은 커뮤니케이션 조건에 따라 크기와 방향이 다르게 나타나 결과물의 질을 높이기 위해서는 보수의 양과 커뮤니케이션 등 두 방식을 모두 고려할 필요가 있었다. 본 연구의 한계점으로는 실험 내용이 클라우드소싱에서 일반적으로 진행되는 작업과 달라 연구 결과를 실제 작업환경에 바로 적용하기는 어렵다는 점이 있었다. 또한 자동화된 메시지 제공 방식을 사용하여 작업자와의 커뮤니케이션이 풍부하지 못했다는 점도 한계로 지적되어, 차후에 보다 풍부한 커뮤니케이션을 제공하는 방안이 필요함이 밝혀졌다.

## 9. 참 고 문 헌

1. 권은주, 김성일 (2003). 성취목표와 피드백의 유형이 내재동기와 지각된 유능감에 미치는 영향. **한국교육학연구**. 41(4). 341-364.

2. 이차연, 박영석.(2006). 성취 목표 지향성이 업무수행 피드백에 미치는 효과. *한국심리학회지: 산업 및 조직*. 19(1). 25-38.
3. 한주희, 고수일 (2005). 결과중심의 피드백 효과에 대한 종단적 연구. *산업경제연구*. 18(5). 2153-2174.
4. Bandura, A. (1989). Self-regulation of motivation and action through internal standards and goal systems. In L. A. Pervin(Eds.). *Goal concepts in personality and social psychology* (pp.19-85). Hillsdale, NJ: Erlbaum.
5. Harackiewicz, J., & Sansone, C. (1991). Goals and intrinsic motivation: You can get there from here. In M. L. Maehr & P. R. Pintrich(Eds.). *Advances in motivation and achievement* (Vol.7, pp.21-49). Greenwich, CT: JAI Press.
6. Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting and task performance*. Englewood cliffs, NJ: Prentice Hall.
7. Akkaya, C., Conrad, A., Wiebe, J., & Mihalcea, R. (2010). Amazon mechanical turk for subjectivity word sense disambiguation. *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk*, 195-203.
8. Bargh, J. A., Gollwitzer, P. M., Lee-Chai, A., Barndollar, K., & Trötschel, R. (2001). The automated will: Nonconscious

- activation and pursuit of behavioral goals. *Journal of Personality and Social Psychology*, 81, 1014–1027.
9. Bernstein, M. S., Little, G., Miller, R. C., Hartmann, B., Ackerman, M. S., Karger, D. R., Crowell, D., & Panovich, K. (2010). Soylent: A Word Processor with a Crowd Inside. *Proceedings of the 23rd annual ACM symposium on User Interface software and technology*, 313–322.
  10. Boner, S. E., Hastie, R., Sprinkle, G., & Young, S. M. (2000). A review of the effects of financial incentives on performance in laboratory tasks: Implications for management accounting. *Journal of Management Accounting Research*, 12, 19–64.
  11. Campion, M. A., & Lord, R. G. (1982). A control systems conceptualization of the goal-setting and changing process. *Organizational Behavior and Human Performance*, 30, 265–287.
  12. Dow, S. P., Kulkarni, A., Bunge, B., Hartmann, B., Nguyen, T., & Klemmer, S. R. (2011). Shepherd the Crowd: Managing and Providing Feedback to Crowd Workers. *Proceedings of CHI EA'11: Extended Abstracts on Human Factors in Computing Systems*, 1669–1674.

13. Epstude, K., & Roese, N. J. (2008). The functional theory of counterfactual thinking. *Personality and Social Psychology Review, 12*, 168–192.
14. Erez, M. (1986). The congruence of goal-setting strategies with sociocultural values and its effect on performance. *Journal of Management, 12*, 585–592.
15. Erez, M., & Kanfer, F. H. (1983). The role of goal congruence in goal setting and task performance. *Academy of Management Review, 8*(3), 454–463.
16. Erez, M., Earley, P. C., & Hulin, C. L. (1985). The impact of participation on goal acceptance and performance: A two step model. *Academy of Management Journal, 28*, 50–66.
17. Finin, T., Murnane, W., Karandikar, A., Keller, N., Martineau, J., & Dredze, M. (2010). Annotating named entities in twitter data with crowdsourcing. *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk*, 80–88.
18. Gallien, T., & Oomen-Early, J. (2008). Personalized Versus Collective Instructor feedback in the Online Courseroom: Does Type of Feedback Affect Student Satisfaction, Academic



Performance and Perceived Connectedness With the Instructor?.  
*International Journal on E-Learning*, 7, 464–476.

19. Harackiewicz, J. M., Abrahams, S., & Wageman, R. (1987). Performance evaluation and intrinsic motivation: The effects of evaluative focus, rewards, and achievement motivation. *Journal of Personality and Social Psychology*, 53, 1015–1023.
20. Kim, S. H., Yun, H., & Yi, J. S. (2012). How to filter out random clickers in a crowdsourcing-based study?. *Proceedings of the 2012 BELIV Workshop: Beyond Time and Errors – Novel Evaluation Methods for Visualization*.
21. Kittur, A., Chi, E.H., & Suh, B. (2008). Crowdsourcing user studies with Mechanical Turk. *Proceedings of the SIGCHI conference on Human factors in computing systems*, 453–456.
22. Kochhar, S., Mazzocchi, S., & Paritosh, P. (2010). The anatomy of a large-scale human computation engine. *Proceedings of the ACM SIGKDD Workshop on Human Computation*, 10–17.
23. Kulik, J. A., & Kulik, C. T. (1988). Timing of feedback and verbal learning. *Review of Educational Research*, 58, 79–97.
24. Latham, G. P., & Seijts, G. H. (1999). The effects of proximal and distal goals on performance on a moderately complex task. *Journal of Organizational Behavior*, 20, 421–429.

25. Latham, G. P., & Steele, T. P. (1983). The motivational effects of participation versus goal setting on performance. *Academy of Management Journal*, 26, 406–417.
26. Locke, E. A. (1968). Toward a theory of task motivation and incentives. *Organizational Behavior and Human Performance*, 3, 157–189.
27. Locke, E. A., & Latham, G. P. (2002). Building a Practically Useful Theory of Goal Setting and Task Motivation: A 35–Year Odyssey. *American Psychologist*, 57, 705–717.
28. Mason, W., & Suri, S. (2012). Conducting behavioral research on Amazon’s Mechanical Turk. *Behavior Research Methods*, 44, 1–23.
29. Mason, W., & Watts, D. J. (2009). Financial incentives and the “performance of crowds”. *ACM SIGKDD Explorations Newsletter*, 11(2), 100–108.
30. Pintrich, P. R. (2000). An Achievement Goal Theory Perspective on Issues in Motivation Terminology, Theory, and Research. *Contemporary Educational Psychology*, 25, 92–104.
31. Roese, N. J. (1994). The functional basis of counterfactual thinking. *Journal of Personality and Social Psychology*, 66, 805–818.

32. Ross, J., Irani, L., Six Silberman, M., Zaldivar, A., & Tomlinson, B. (2010). Who are the crowdworkers?: shifting demographics in mechanical turk. *Proceedings of CHI'10 Extended Abstracts on Human Factors in Computing Systems*, 2863–2872.
33. Shaw, A. D., Horton, J. J., & Chen, D. L. (2011). Designing incentives for inexperienced human raters. *Proceedings of the ACM 2011 conference on Computer supported cooperative work*, 275–284.
34. Smallman, R., & Roese, N. J. (2009). Counterfactual thinking facilitates behavioral intentions. *Journal of Experimental Social Psychology*, 45, 845–452.
35. Snow, R., O'Connor, B., Jarafsky, D., & Y. Ng, A. (2008). Cheap and fast – but is it good?: evaluation non-expert annotations for natural language tasks. *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, 254–263.
36. Ullmer, B., & Ishii, H. (1997). The metaDESK: models and prototypes for tangible user interfaces. *In Proceedings of the 10th annual ACM symposium on User interface software and technology*, 223–232.

37. Whitehill, J., Wu, T. F., Bergsma, J., Movellan, J. R., & Ruvolo, P. L. (2009). Whose vote should count more: Optimal integration of labels from labelers of unknown expertise. *Advances in neural information processing systems*, 2035–2043.

## ABSTRACT

# The Effectiveness of Feedback and Goal on The Quality of Crowdsourcing Result

Lim, Jae-Eun

Department of Communication

The Graduate School

Seoul National University

The quality improvement has been an important issue in crowdsourcing management. Prior studies have tried to improve the quality by amount or scheme of payment, filtering results, or voting but the ways increased quality only in restricted conditions. In this paper, we suggested communication as a new way to manage quality. From prior studies, types of work-related communication were feedback and goal and each type consisted of 4 categories. Before the main experiment, prior study was performed to find out the effect of communication in crowdsourcing. When message was given that someone would evaluate results, workers spend more time but the quality was under expectation. The finding was that

work-related communication was needed to be specific for better result. With this insight, the main experiment was conducted from May 26 to June 10, 2015. Each participant was assigned to one of 8 communication styles and one of 2 payment schemes were set to find the effect of payment on quality management. Dependent variables were final answer's length, work time, and cost. As results, negative feedback was more effective to improve the quality than positive feedback as it reduced the length, time, and cost but the positive one didn't. Evaluative feedback was found to improve the quality better than descriptive feedback by reducing length and cost. In terms of goals, proximal & distal goal reduced the length and time more than distal goal, and the amount of reducing length was similar so proximal & distal goal was proved to be more effective for improving quality. Achievement goal was shown to improve quality better than target goal. We found that the most appropriate communication in crowdsourcing environment need to be negative and evaluative, the goal should be specific with proximal & distal goals, and context is required to improve the work quality. Negative feedback, evaluative feedback and proximal & distal goal were effective because they can help workers judge their results clearly. And workers can be engaged more when they understand the

reason and context of works. Also from the results, the amount of payment affected the quality as prior researches insisted but the effect was seen to be different by communication conditions. With these findings, this research found the possibility of communication as a solution for quality management in crowdsourcing.

**Keywords:** Crowdsourcing, result quality, feedback, goal, Amazon's Mechanical Turk

*Student Number:* 2013-22844